DOCUMENT RESUME

ED 405 183 SE 057 380

AUTHOR	Brunelli, Lina, Ed.; Cicchitelli, Giuseppe, Ed.
TITLE	Proceedings of the Scientific Meeting of the International Association for Statistical Education (1st, Perugia, Italy, August, 1993).
INSTITUTION	International Association for Statistical Education.
PUB DATE	94
NOTE	449p.; Financial support provided by UNESCO Italian National Council for Research, and University of Perugia.
AVAILABLE FROM	University of Perugia, via A. Pascoli, 06100 Perugia, Italy.
PUB TYPE	Books (010) -- Collected Works - Conference Proceedings (021)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Computer Software; Computer Uses in Education; *Data Analysis; Developing Nations; Educational Innovation; *Educational Research; Elementary Secondary Education; Engineering Education; Faculty Development; Foreign Countries; Government Employees; Higher Education; Information Technology; Learning Theories; Mathematics Education; Models; *Probability; Professional Training; *Statistics; Student Surveys; Teaching Methods; *Telecommunications; Television
IDENTIFIERS	Biostatistics; International Association for Statistical Educ

ABSTRACT

This volume contains papers on computing and television as well as on the rapidly growing area of electronic communication among teachers and between teachers and students. Data analysis is now an important strand in school curricula in many nations. Many of the contributors discuss statistics in the schools from various perspectives. University education is not neglected, nor is the issue of training government statistical staff. This proceedings is divided into eight sections according to general topic. The sections are: (1) "Statistical Education at School Level"; (2) Teaching Probability and Statistics at University Level"; (3) "Computers, Video and Other Tools in the Teaching of Probability and Statistics"; (4) "Education Programs and Training in Statistics"; (5) Issues in the Teaching of Probability and Statistics"; (6) "The IASE [International Association for Statistical Education] and Problems of Statistical Education in Developing Countries"; (7) "Updating Teaching Methods in Probability and Statistics"; and (8) abstracts from the poster sessions. Contains a list of participants. (PVD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

405-183

PROCEEDINGS
OF THE
FIRST SCIENTIFIC MEETING

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

G. Cecchitelli

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SE057380.

BEST COPY AVAILABLE

INTERNATIONAL ASSOCIATION FOR STATISTICAL EDUCATION

PROCEEDINGS
OF THE
FIRST SCIENTIFIC MEETING

Edited by

Lina Brunelli and Giuseppe Cicchitelli

University of Perugia, 23-24 August 1993

© Università di Perugia, 1994

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

Editors: Lina Brunelli and Giuseppe Cicchitelli
Dipartimento di Scienze Statistiche
Università di Perugia
Via A. Pascoli
06100 Perugia (Italy)

Printed by Centro Stampa dell'Università di Perugia

The financial support of UNESCO, Italian National Council for Research (CNR) and University of Perugia is gratefully acknowledged.

PREFACE

This volume contains all the invited and contributed papers presented at the first scientific meeting of the International Association for Statistical Education (IASE), held in Perugia in August 23 and 24, 1993. The papers have been collected together under appropriate topic headings, regardless of the order in which they appeared in the scientific programme.

A report about a session devoted to Italian school teachers on "updating teaching methods", and the abstracts of the papers presented at the poster session are also included.

The scientific programme of the meeting was arranged by the IASE interim executive headed by Professor David Vere-Jones with the cooperation of the local organising committee headed by Professor Giuseppe Cicchitelli. The full membership of the two committees was:

IASE interim executive: D. Vere-Jones (Chairman), D. Berze, K. Bryson, Y. Escoufier, A. Hawkins, M.J. Laurent-Duhamel, J. Ntozi, L. Pereira-Mendoza, M. Regier, R. Scheaffer, K. Vännman.

Local organising committee: G. Cicchitelli (Chairman), L. Brunelli, G. Galmacci, R. Giaimo, M. G. Ottaviani, M. A. Pannone, S. Rigatti Luchini.

We are indebted to Professor Vere-Jones, not only for his authoritative contribution to the arrangement of the programme, but also for his many kind suggestions in relation to the organisational aspects.

We are grateful to UNESCO, to the National Council for Research, to the University of Perugia and to other local Institutions for their financial support.

We express our gratitude to Kenneth Bryson, Anne Hawkins, David Moore, Brian Phillips, Mary Regier, Richard Scheaffer and David Vere-Jones for their invaluable help in editing the volume.

Perugia, May 1994

Lina Brunelli
Giuseppe Cicchitelli

Contents

III Preface

XI Introduction

STATISTICAL EDUCATION AT SCHOOL LEVEL

3 Teaching statistics at school level in some European countries
P. Holmes

13 The teaching of probability and statistics in schools: issues for interdisciplinary reflection
E. Lombardo, C. Rossi and A. Zuliani

27 The classic definition of probability and fractions: difficulties and possibilities at the primary school level
C. Caredda and M. R. Puxeddu

37 Teach-Stat: a professional development program for elementary teachers (grades k-6) in North Carolina, USA
D. Frierson, S. Friel, S. Berenson, G. Bright and C. Tremblay

47 On the psychopedagogical aspects of teaching statistics at the primary school
M. P. Perelli D'Argenzio

53 Stat-Maps
M. Perry

63 Statistics: a motivator for secondary school mathematics
R. L. Scheaffer

TEACHING PROBABILITY AND STATISTICS AT UNIVERSITY LEVEL

73 What engineers need to learn about statistics
G. Box

89 Simulation and the student engineer
L. Råde

97 Off-campus statistical experiences for undergraduates
C. J. Blumberg

105 The effect of context on the teaching of statistics at first year university level
M. Clark

115 The educational tracks of "the making statistics more effective in schools of business" conferences
J. D. McKenzie, Jr.

123 Statistical education in developing countries: the problem of 'marketing' university graduates
H. Nikolov, E. Keogh and K. Stielau

133 University level statistical education in Hungary
J. Reiczigel

141 An investigation into the performance of first year students in mathematics and statistics at a South African university
K. Stielau

149 Teaching biostatistics to medical students
R. van Strik

155 Education in statistics at the Higher Institute of Chemical Technology - Sofia
K. Velev, I. Vuchkov and H. Iontchev

COMPUTERS, VIDEO AND OTHER TOOLS IN THE TEACHING OF PROBABILITY AND STATISTICS

161 The role of electronic communication in statistics education
T. Arnold

173 Cognitive technologies for statistics education: relating the perspective of tools for learning and of tools for doing statistics
R. Biehler

191 Using computers in teaching statistics - Is it worthwhile?
A. Blejec

199 Understanding probability and statistical inference through resampling
C. Konold

213 The uses of video in teaching statistics
D. S. Moore

221 Educational hypertext in the teaching of statistics: a possible use for an instruction module for the analysis of the relations between two variables
F. Camillo and M. Pedroni

229 Statree: an expert system for choosing suitable statistical data processing techniques
C. Capiluppi and L. Fabbris

237 Designing curricular software for conceptualizing statistics
S. Cohen, G. Smith, R. Chechile and R. Cook

247 The use of a computer based advisory system as a teaching aid in statistics
J. Griffin and M. Rouncefield

253 A computer aided software for teaching statistics
M. Janvier, G. Saporta and E. Verdoire

255 A role for computer intensive methods in introducing statistical inference
P. Jones, K. Lipson and B. Phillips

265 Chemometrical package for PC
J. Militký, M. Meloun and K. Kupka

273 A computer project for teaching statistics to technicians and factory workers
A. M. Morin

277 Teaching data analysis with the help of Survo
J. Puranen

EDUCATION PROGRAMMES AND TRAINING IN STATISTICS

285 The role of consultancy in university education and professional training in statistics
V. Barnett

299 Three decades of training government statistical staff in developing countries: the African experience
J. P. M. Ntozi and G. W. Kibirige

303 The Eurostat project "Training of European Statisticians" (TES): experience of the first three years and some reflections on the future
R. Teekens

315 Teaching biostatistical concepts and methodology to medical professionals: various alternatives
S. I. Bangdiwala

321 Teaching statistics to production personnel
A. C. García

327 ISEC, Calcutta: The pioner centre for statistical training of government sponsored statisticians from developing countries of the third world, mainly from Asia and the Far East
A. B. Raha

333 Problems of statistical education in Croatia
A. Rozga

337 Statistical education programme in Pakistan
S. A. Starkings

ISSUES IN THE TEACHING OF PROBABILITY AND STATISTICS

345 Student views on effective and ineffective learning in graduate level applied statistics courses
G. M. Belli

355 Intuitive strategies for teaching statistics
M. G. Borovcnik

367 Secret codes in statistical education
E. Çakiroğlu and T. Nemetz

377 Proportions, probability and other matters
F. Jolliffe

385 Teaching the essentials of elementary probability theory or how not to calculate the mean of the negative hypergeometric distribution
P. Nüesch

395 Statistical simulations
M. Rouncefield

401 Teaching statistics: making it memorable
E. R. Sowey

409 What is the main goal of statistical education?
V. V. Shvyrkov

417 Learning methods involving communication by students
J. Taffe

THE IASE AND PROBLEMS OF STATISTICAL EDUCATION IN DEVELOPING COUNTRIES

429 The IASE and problems of statistical education in developing countries
D. Vere-Jones

UPDATING TEACHING METHODS IN PROBABILITY AND STATISTICS
Session devoted to Italian school teachers

445 Updating teaching methods in probability and statistics
M. G. Ottaviani

POSTER SESSION

453 Abstracts:
A. Bartkowiak
M. Davì
Y. Ersoy and T. Nemetz
R. Giaimo and A. M. Milito
S. N. Habibullah
R. Hernandez
A. Mascelloni
G. Moncecchi and M. P. Perelli D'Argenzio
J. Panaretos
P. Ranzani and M. De Poli
J. Taffe
E. Xekalaki
J. Yang and N. Li

465 LIST OF PARTICIPANTS

INTRODUCTION

At its 48*th* session in Cairo in 1991, the International Statistical Institute approved the creation of a new section, or affiliated society, the International Association for Statistical Education (IASE). The new society held its first scientific meeting in Perugia, Italy, on August 23 and 24, 1993. More than 150 persons from 36 countries participated in the Perugia meeting.

The papers in this volume suggest the variety and importance of the issues facing statistical education as the IASE begins its existence. What is the place of technology in teaching? Readers will find here papers not only on computing and video, but on the rapidly growing area of electronic communication among teachers and between teachers and students. Data analysis is now an important strand in school curricula in many nations. A number of authors discuss statistics in the schools from various perspectives. Note in particular Peter Holmes's comparison of practices within the EEC, which is particularly appropriate for IASE, because it moves beyond the experiences of a single national system.

University education is not neglected, nor is the issue of training government statistical staff.

The IASE, like the ISI itself, is acutely aware of the special problems facing statistical education in developing countries and in the transition countries of eastern Europe. I am pleased that invited speakers from Slovenia, Uganda, and Zimbabwe were at Perugia, and I hope that readers will carefully ponder David Vere-Jones's eloquent essay on "The IASE and Problems of Statistical Education in Developing Countries". David Vere-Jones was the interim president of IASE while the society was in its formative stages. We owe him and the other members of the interim executive committee our gratitude for launching IASE. We are also indebted to Giuseppe Cicchitelli and the other members of the Perugia organizing committee for arranging and hosting so excellent a meeting, and for attracting financial support for the meeting at a time when IASE had few resources of its own.

The open forum on future directions for IASE that concluded the meeting allowed participants to suggest priorities for the new society. Among many ideas, I noted a repeated emphasis on finding ways to allow people concerned with statistical education to more easily locate and communicate with each other. Membership in IASE is of course a way of announcing a commitment to statistics education; we hope to offer members innovative ways to reach each other electronically and otherwise.

The IASE aims to be an effective international voice for strengthening statistical education at every level. It is, I believe, the first society having this as its primary purpose. The Perugia meeting and the papers appearing in this volume are a step in that direction.

David S. Moore
President, IASE

STATISTICAL EDUCATION AT SCHOOL LEVEL

Invited papers	Peter Holmes Enzo Lombardo *et al.*
Contributed papers	Carla Caredda *et al.* Dargan Frierson *et al.* Maria Pia Perelli D'Argenzio Mike Perry Richard L. Scheaffer

TEACHING STATISTICS AT SCHOOL LEVEL IN SOME EUROPEAN COUNTRIES

Peter Holmes
Centre for Statistical Education, University of Sheffield
Sheffield, S3 7RH, England

1. Introduction

In 1977, the International Statistical Institute set up a task force on the teaching of statistics at the school level (TOTSAS) under the chairmanship of Professor Vic Barnett. Amongst other things, this committee published a book entitled "Teaching Statistics in Schools throughout the World". The section on Europe included reports on England and Wales, Scotland, the Federal Republic of Germany, France, Hungary, Italy and Sweden. This paper reports on some of the changes in all of these countries except Hungary, and also includes a reference to the current position in Spain.

The original survey showed a wide range of approaches to the teaching of statistics between the countries. There was disagreement on the amount of statistics, the nature of the appropriate material, the age at which it should be introduced, and the balance between probability and statistics.

Barnett reported on the situation in England and Wales. In the primary schools there were no national guidelines and, although some good work on data collection and pictorial presentation could be found in some schools, in others there was very little. For pupils aged 11-16 there was little statistics in the standard mathematics course but there were optional courses in probability and statistics that had quite full syllabuses. He reported on a major curriculum development project, the Schools Council Project on Statistical Education, which produced teaching material with the aim of introducing statistics for citizenship, stressing the practical relevance of statistics by introducing concepts and techniques almost entirely through the medium of detailed real-life problems, chosen from a wide range of application areas. Barnett also reported on the setting up of a Centre for Statistical Education at the University of Sheffield.

Dinges, from Germany, reported that there was no uniform picture. He noted that the views of what stochastics should be differed between states, and that states varied from those with centrally controlled written examinations to those where individual schools had more freedom. Stochastics was acknowledged in both basic and more intense courses within mathematics. Although probability could vary from the intuitive to Kolmogorov's axioms, the tendency was to formalise probability to events in the sample space and formal manipulation rules. He expressed a desire to see a change to the position where applications should not merely be seen as illustrations to enhance motivation.

From France, Hennequin reported that there was no work on probability for pupils younger than 14. He noted that teaching statistics remained formal because of lack of time and detected an over-emphasis on formal probability, involving questions on balls in urns. No reference was made to any statistical work in primary schools.

In Italy, Zuliani noted that the syllabus for 11-13 year olds included *ordering and correlating data* and *drawing attention to the differences between certain and probable.* Of six theses in the mathematics syllabus, one concerned statistics. There was a more complex picture for the upper secondary schools (age 14+, courses up to 5 years long). Vocational schools included some stochastics, mostly probability theory and description. There was some simple descriptive statistics, calculating simple statistical measures. Again no reference was made to any work in primary schools.

Ekenstam reported that in Sweden courses for grades 4-6 (age 10-13) included the collection of statistical data, tables and graphs. The work for grades 7-9 included frequency tables, graphs, mean, median, some measure of spread and probability. He found directions that suggested grade 1-3 children should gather and report information in pictorial form.

That was the position at the end of the 1970's. How does it differ today?

2. Why do we teach statistics?

It is clear that there are two different traditions meeting here. What we choose to teach will depend on our philosophy and definition of statistics.

One tradition has a very broad view of statistics. It sees statistics as part of society, including uses made of data in all walks of life. Since, on this definition, it is a numerate subject, it is up to mathematics teachers to be responsible for this teaching - but they will not be the only teachers to be so involved. Statistics will also occur in the teaching of science, geography, economics, etc. This tradition is summed up in the following quotation from the British Cockcroft Report from (Cockcroft, 1981).

> §781. Statistical numeracy requires a feel for numbers, an appreciation of appropriate levels of accuracy, the making of sensible estimates, a common-sense approach to the use of data in supporting an argument, the awareness of the variety of interpretation of figures, and a judicious understanding of widely used

concepts such as means and percentages. All these are part of everyday living. Good statistics can encourage pupils to think in these ways.

This philosophy of statistics will mean that at least lip service is paid to the place of statistics in other areas of the curriculum. Again from the Cockcroft report:

> §776. Statistics is essentially a practical subject and its study should be based on the collection of data, wherever possible by the students themselves. It should consider the kinds of data it is appropriate to collect, the reasons for collecting the data and the problems of doing so, the ways in which the data may legitimately be manipulated and the kinds of inference which may be drawn. Work in subjects such as biological science, geography and economics can therefore contribute to the learning and understanding of statistics.

The non-statutory guidelines of the National Curriculum in Mathematics for England and Wales (Department of Education and Science, 1989) refer to having a mathematics policy for the whole school and using data from other subjects. In Spain, the Andalucian expansion of the Ministry of Science and Education syllabus refers to the place of probability and statistics in our world culture. In Italy, the *programma didattico* for *scuola elementare* stresses the use of mathematics in subjects to help scientific understanding of problems and a logical aproach to solutions. It also stresses the practical uses of mathematics - which by implication includes the sections on probability and statistics. The scheme for the upper secondary schools in Italy refers to the importance of statistics in science and real life in general.

The second tradition that can be detected behind the proposed syllabuses is that of mathematical statistics or stochastics. In Germany *stochastics* also incorporates applied probability, including risk analysis and decision making under uncertainty. In this tradition, statistics is seen as a subset of mathematics (including applied mathematics) and the main concern is with modelling uncertainty and developing theoretical models. There is a much greater emphasis on probability and distributions, and much less on data and data collection. This attitude can be seen behind goals such as those described by Nemetz (in Barnett, 1981, p. 97) for Hungary:

> a. Make the pupils acquainted with the notion of randomness. Make them able to view stochastically the notion of random events.
>
> d. Get the children acquainted with basic laws of probability theory and the main ideas of mathematical statistics.

e. Show them how to construct stochastic models in real life problems. (goal *b* was on evaluating statistical data and the limitations of stochastic arguments, and goal *c* was on the stability of relative frequencies).

The ASU group from France (Ambroise *et al.*, 1991) refer to this tradition when they criticise the positioning of the phrase *organization et gestion des données* immediately next to *fonctions*. They also refer to textbooks where statistics is only an excuse for probability. Morin (1991) writes of the strong influence of mathematics and pure mathematicians.

At its worst this tradition can lead to sterile theoretical models and much work in pure mathematics, with no real thought about applications. At its best it can make pupils think deeply about the structure of a real life problem and see how to work to its solution using probability theory. At its best it can also sharpen up concepts in statistical inference.

Fortunately, there does seem to be some coming together of these two strands in the countries studied, although the actual balance reflects society and the history of mathematics teaching in each country. As an example, two successive aims in the Swedish syllabus are:

> to develop some knowledge and skills in arithmetic, geometry and statistics and insights in the concepts of probability and functions, and
>
> to develop basic statistical concepts and methods for collecting and handling data and using tables and diagrams to describe and compare important properties of statistical information.

Even so, fine statements such as these need to be matched against what occurs in the classroom. The reality can be very patchy, particularly where teachers are insecure in their own statistical knowledge and teaching ability. They may avoid teaching statistics if they can and, even where they do teach it, they may find it difficult to meet such high aims.

3. What statistics should we teach?

It is interesting to note that in many countries the word *statistics* did not convey what the course designers wanted it to include. In some countries it was associated with too narrow a view (as might be implied by the word *stochastics*). In others, it was associated too much with official statistics and the amassing of numerical information with no analysis or theoretical inference.

In the United Kingdom, in an attempt to solve this problem, the content is put under the heading of *Handling Data* while French syllabuses refer to *Data Organisation.* In Germany, you may find some of this work under *Sachrechnen* (Calculating in a Context). Such phrases appear to have been chosen to emphasise an active approach to data and the idea of using real data to obtain meaningful insights and draw meaningful inferences. They are a good counterbalance to the tendency to teach techniques for their own sake, based on fictitious data.

There is a common core of techniques and concepts at the heart of the syllabuses from all the countries considered. These include summarising the data in pictorial form such as bar charts or histograms; using and calculating measures of central position such as the mean and the median; having some idea of variability, which may be measured by interquartile range or even standard deviation; and probability, as a measure of uncertainty and the basic rules of manipulating probabilities (at least for independent and for mutually exclusive events).

There is not a common pattern with reference to the inclusion of EDA techniques such as stem and leaf or box plots, nor to the idea of exploratory data analysis as such. Neither is there a common reference to sampling and the fundamental statistical idea of using a sample to draw inferences about a population. There is no common reference to the distinction between a survey and an experiment, nor to the fundamental idea of randomisation in experimental design. There is no common attempt to try to get pupils to appreciate the importance of statistics in society nor to think statistically when making decisions. Although not one of these things is common, elements of each occur in some of the syllabuses. Other courses include reference to the use of computers, databases, spreadsheets and statistical packages.

The reports from France seem to be rather pessimistic. The ASU, after noting that primary pupils could collect data, put them into tables, and give graphical representations, write that on the face of it the situation seems good but that the reality is different. Apparently, most teachers do not teach that part of the programme. Morin (1991) reports that teachers are supposed to deal with data organisation in the last year of primary school, but whether this is actually done depends on the individual teacher's attitude to statistics. The BEPC does include a course in basic statistics for all pupils aged 11 to 16, but it is estimated that only one third of teachers teach statistics inside the mathematics classes. In a personal communication, Morin reports that the BEPC test in *Geography* is full of statistics. Students must draw histograms, comment on statistical tables and handle conditional distributions. In short, in geography the students handle more statistics than in mathematics. Reports from Germany also indicate a growing use of computer supported data analysis in subjects other than mathematics.

The approach to teaching probability varies immensely. To some extent it depends on the age at which the ideas are first introduced. In England and Wales the approach is through events that may or may not happen, with the idea of *more* or *less likely* being hardened up to a rank order that can be expressed on a 0 to 1 scale. *Equally likely outcomes* are encountered as a special case, and *relative frequencies* are introduced after equally likely events. The Italian primary school syllabus recommends that the probability should be introduced through real life and through games involving judgements about *perhaps, certain* or *possibly.* Later it recommends using games to learn probability. In other countries, the equally likely approach predominates and is much more closely linked with combinatorics. In Spain, the Andalucian expansion of the Ministry of Science and Education syllabus includes combinatorics and sample spaces leading to conditional probability.

4. When should we teach probability and statistics?

On this point, there is a wide range of views. Most states in Germany include some work in probability and statistics in mathematics before the age of 16 (although there are some exceptions where none is included). All states include stochastics for grades 11 to 13 (the next age goup). The most common practice is for some such work to be included at least in the syllabuses for pupils aged 10+. For some countries the start is made much earlier.

In England and Wales (Department of Education and Science, 1991), work on sorting (leading to data representation) starts at the age of 5. Subjective use of probability words such as *uncertain, possibly* etc. leading to ranking probabilities begins soon afterwards. In Scotland (Scottish Office Education Department, 1991), data representation begins in the same way, but there is no work on probability until after the age of 14. Although France allows such work in primary schools, apparently not as much work happens as our French colleagues would like. A more positive start is made at the secondary level. In Spain (Godino, 1992), primary school pupils are taught to use elementary techniques of data recognition to obtain information, and to represent it in graphical and numerical form. Some elementary probability is also introduced. States in Germany tend to start at the secondary level. Sweden starts in the early years of the primary school, so that at the end of five year's schooling Swedish pupils should be able to interpret data in tables or diagrams and to use the concepts of median and mean. Italy has a detailed programme for primary schools including probability and statistics. These pupils are expected to be able to recognise data sufficiency and see contradictions in data.

Wherever recommendations are made about how statistics should be taught they are always along the lines of doing practical work, using real data, and teaching in real life contexts.

5. Other considerations

5a. *Central Direction*

There is a wide range in the amount of central direction of the teaching content. At one extreme is Germany where the decisions are left to the individual states, and even then there may be only outlines which schools or individual teachers have to interpret. At the other extreme is the United Kingdom. From being one of the most diverse systems at the time Barnett was writing in 1982, it has now become the most centralised of all the countries being considered here. The National Curriculum in Mathematics for England and Wales (there is a separate but similar document for Scotland) applies in detail to all the schools in the country. It gives a highly detailed syllabus, including work in probability and statistics, split into ten levels that are to be covered in the age range 5 to 16. There are national tests at ages 7, 11, 14 and 16.

Spain is much more typical. There, a general syllabus is laid down nationally and local areas put in their own detail. Italy and Sweden have similar systems

5b. *Testing*

It has not been possible to obtain examples of the different ways used to test students' understanding of probability and statistics. This, though, is a most important aspect of teaching. The nature of the test is probably the single biggest factor in determining how the subject is taught. To borrow, more or less, a phrase from word-processing "What You Test Is What You Get".

If the teacher sees that the assessment emphasises something then (s)he will rightly teach the students to be able to pass this assessment. If the assessment only tests the trivial then that is what will be taught. More fundamental understanding will then tend to be ignored. This is a matter for concern in the United Kingdom.

The ASU group in France put the matter well when they reported that there was a lack of statistics questions in examinations: questions test calculating ability and memory, but not a critical spirit.

6. Conclusion

There does seem to be a coming together in Europe concerning what should be taught in probability and statistics at the school level when compared with 10 to 15 years ago. There is, though, a lot that we can still learn from each other and there are many ways in which we could be of mutual help.

Bibliography

Ambroise D. *et al.* (1991), *L'enseignement de la statistique en France*, Association pour la Statistique et ses Utilisations.

Barnett V. (ed.) (1982), *Teaching Statistics in Schools throughout the World*, International Statistical Institute, Voorburg.

Cockcroft W.H. (1981), *Report on the Committee of Enquiry into the Teaching of Mathematics*, HMSO.

Department of Education and Science (1989), *Mathematics. Non-Statutory Guidelines*, HMSO.

Department of Education and Science (1991), *Mathematics in the National Curriculum*, HMSO.

Godino J. (1992), Teaching Statistics in Primary and Secondary Schools in Spain, in the ISI newsletter insert in *Teaching Statistics*, 14(2), Teaching Statistics Trust.

Hawkins A. S., Jolliffe F. and Glickman L. (1992), *Teaching Statistical Concepts*, Longman, London.

Morin A. (1991), Teaching Statistics at the Secondary School Level in France, in Vere Jones D. (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg.

Scottish Office Education Department (1991), *Mathematics 5-14*, HMSO.

Syllabuses in mathematics were obtained from the following countries:

England and Wales
Germany - several of the federal states
Italy - national syllabus and the expansion used in Umbria
Scotland
Spain - the national syllabus and the Andalucian expansion
Sweden
Finland

Acknowledgements

I would like to thank the following for providing information and generally helping me with the preparation of this paper:

C. Batenero	University of Granada
R. Biehler	University of Bielefeld
L. Brunelli	University of Perugia
G. Cicchitelli	University of Perugia
A. Dunkels	University of Luleå
J. Godino	University of Granada
R. Holmes	University of Hull
B. Johansson	University of Gothenburg
A. Morin	University of Beaulieu, Rennes
M. Pannone	University of Perugia
L. Råde	University of Gothenburg

THE TEACHING OF PROBABILITY AND STATISTICS IN SCHOOLS: ISSUES FOR INTERDISCIPLINARY REFLECTION

Enzo Lombardo
Dipartimento di Studi Geoeconomici, Statistici e Storici
Università di Roma "La Sapienza"
Via Castro Laurenziano 9, 00161 Roma, Italy

Carla Rossi
Dipartimento di Matematica, Università di Roma "Tor Vergata"
Via Orazio Raimondo, 00173 Roma, Italy

Alberto Zuliani
Dipartimento di Studi Geoeconomici, Statistici e Storici
Università di Roma "La Sapienza"
Via Castro Laurenziano 9, 00161 Roma, Italy

1. Introduction

In the following discussion several issues are presented for interdisciplinary reflection on the teaching of Probability and Statistics in preparatory school. They are set out according to a scheme which links this particular subject, with, in the first place, other disciplines – mathematics, physical-natural sciences, humanities and social sciences; secondly, with the variables of individual learning and modes of scholastic organization; and finally, with environmental variables (see scheme 1). The scheme is organized around problems and disciplines involved in the didactic complex and incorporates many features of the learning-teaching process.

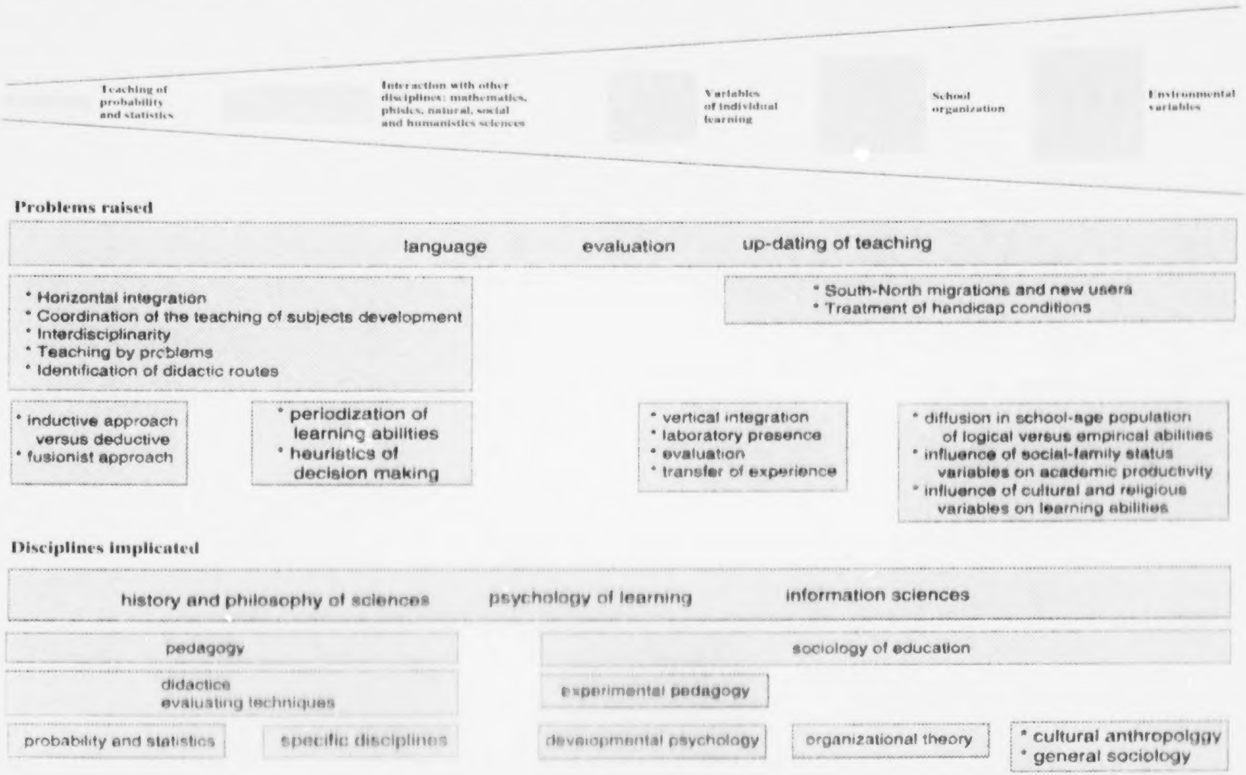
Some of the issues originate in the Italian reality; they may also be valid in other national contexts.

2. The conditioning of tradition (the Italian case)

Italian culture has strong humanistic traditions. This involves:

- development of linguistic abilities in an instrumental sense for the reading of "classics", to the detriment of the development of technical-scientific languages, contrary to what generally occurs in other countries (for example in Anglo-Saxon ones);

Scheme 1. *Probability and statistics in schools from an interdisciplinary viewpoint*


Teaching of probability and statistics
Interaction with other disciplines: mathematics, phisics, natural, social and humanistics sciences
Variables of individual learning
School organization
Environmental variables
Problems raised
language
evaluation
up-dating of teaching
* Horizontal integration
* Coordination of the teaching of subjects development
* Interdisciplinarity
* Teaching by problems
* Identification of didactic routes
* South-North migrations and new users
* Treatment of handicap conditions
* inductive approach versus deductive
* fusionist approach
* periodization of learning abilities
* heuristics of decision making
* vertical integration
* laboratory presence
* evaluation
* transfer of experience
* diffusion in school-age population of logical versus empirical abilities
* influence of social-family status variables on academic productivity
* influence of cultural and religious variables on learning abilities
Disciplines implicated
history and philosophy of sciences
psychology of learning
information sciences
pedagogy
sociology of education
didactics
evaluating techniques
experimental pedagogy
probability and statistics
specific disciplines
developmental psychology
organizational theory
* cultural anthropolgy
* general sociology

BEST COPY AVAILABLE

- greater depth, in the study of the development of thought, of those parts "close" to logical-deductive schemes (in the study of the History of philosophy one spends more time on Descartes, Kant, Hegel and up to Croce, than on Galileo, English empirism, pragmatism).

On the whole, the didactic courses peculiar to each discipline develop abilities and aptitudes for deductive processes, rather than inductive processes. Even the teaching of physics (and other experimental sciences) is closer to this general scheme than to the Galileian "method". It is done through the study of handbooks rather than by analyzing experimental results, even if simple ones. Often, teachers lament the scarcity of adequate structures. Actually, such scarcity represents a convenient *alibi* for the perpetuation of a bookish approach in experimental sciences also. It follows that it is almost never emphasized that the "rules" used to predict physical "facts" are the result of elaborations and calculations made on the basis of mathematical "models", suggested by "empirical" observations; that observations essential to interpreting the phenomenon examined are made in different contexts, from which features are detected which are considered relevant to the solution of the problem being studied. The planetary model of the atom is no more inadequate than others to explain certain phenomena, but it is not fit to explain all phenomena.

The usual courses in academic teaching are aimed at achieving "definitive truths" rather than at a "method" which, step by step, permits one to arrive at the interpretation of different phenomena with a "sufficient degree of approximation", as Guido Castelnuovo already reminded us at the begining of this century. The approximation concept itself is completely extraneous to the traditional culture and not even the introduction of computers, which necessarily produce results with a limited precision has led to greater awareness of its role, either in a numerical sense or in the sense of a model. The inductive method necessarily involves "approximation" and a "provisional" approach to the results achieved, in addition to a certain subjectivity in interpretation. All these concepts and abilities are not only neglected in traditional Italian teaching, but, in some ways, consciously rejected. Any "innovative" proposal must deal with this generally widespread situation, which also produces teachers. These teachers should be the creators of possible modifications in modalities, not only in relation to content, but also in facilitating the development of critical abilities, which is one of the declared goals of the school course.

3. Teaching by problems, teaching by models

The situation described has led to the affirmation of a methodology of teaching scientific disciplines, in particular of mathematics, by paradigms rather than by problems. One studies Euclidean geometry and calculates all the possible areas of plane figures, however complicated, but one does not talk about the fact that really it is a question of applying a model, that this is just an approximation useful for the study of property (for example, of allotments), because actually the earth's surface is spherical and within spherical geometry the axioms and assumptions of Euclidean geometry are not applicable. Once again the search for absolute truths overwhelms the awareness of approximation. This recognition, which sometimes occurs in relation to non-Euclidean geometries, generally appears as something strange rather than as something necessary for the interpretation of certain "facts".

The different mathematical models of space are never analyzed in a comparative way, yet non-Euclidean patterns are generally introduced in a way that subordinates them hierarchically to the Euclidean pattern.

It is considered preferable to reduce the real world to a "flat" world (without even mentioning it), whose historic necessity and utility are bound to the technically limiting instruments which are traditionally used in the study of geometry: (flat) blackboards and (flat) book and notebook pages. Here also, teaching has not taken advantage of the introduction of computers which would not only permit the use of realistic spherical images, but could also explain, for example, why the air-routes from Europe to America do not follow the parallels, as would seem natural from looking at an atlas (flat image), but follow geodesic lines. Analogously, one prefers to reduce the world to a deterministic world rather than to speak of approximations related to forecasts under conditions of uncertainty, and so necessarily based on probabilistic reasoning. Thus, one always prefers to remain anchored to ideal schemes, very far from the reality which one would like to represent.

An extreme feature of this kind of teaching by hypothetical-deductive schemes appears to be the same in the teaching of Probability. This is often inserted in Analyses and is considered as a particular application of analytic calculations (combinatorial calculation or others) rather than as an autonomous sector of mathematics - one which principally, but not exclusively, uses analysis as a tool, but finds its basis in problems which cannot be confronted within a simple logical-deductive model.

Instead, the most adequate scheme for the teaching of Probability and Statistics is of an inductive type:

SIGNIFICANT PROBLEM --> RESOLUTIVE MODEL --> RESULTS (FORECASTS) --> COMPARISON WITH DATA --> POSSIBLE MODIFICATION OF THE PREVIOUS MODEL etc.

It is important both to resolve the same (interesting) problem with many theoretical models, and to apply the same model to many significant problems. Above all, one should avoid using "recipe-books", and transmitting the (wrong) idea of a one-to-one correspondence between problems and models. That is, one must aim at developing the ability to "disentangle" new problems, not inserted within pre-constituted schemes, by using competences acquired through the solution of other problems.

4. Fusionism and interdisciplinarity

New teaching methods, better suited to developing independence in confronting new problems and oriented towards producing solutions based on experience with analogous cases, should be based on hypotheses of fusionism and work on interdisciplinary projects. From interdisciplinary work it is possible to map out interesting and significant problems to be confronted, and also to use the new capabilities given by computers (analysis of complex data, simulation).

The choice, in certain fields of application, of significant problems which are interesting because of the methods used in their solution, facilitates the clarification of both methodological notions and those related to application. Probability and Statistics represent the principle means for dealing with significant applications, and they permit the complete realization of the Galileian scheme:

OBSERVATION --> CONSTRUCTION OF A MODEL --> CALCULATION OF FORECASTS --> DETECTION OF NEW ADEQUATE DATA --> COMPARISON WITH THE FORECASTS.

This scheme also suggests a temporal order in introducing arguments: methods for data description, probabilistic models, inference. One can propose several prototypes of this kind of course, which are both "historical" and current and can help overcome the lack of consolidated experiences. Analyses of data from the field of (classical) physics is particularly instructive because such data correspond to essentially deterministic models in which the only fluctuations are due to errors of measurement. By contrast, analyses of data from biological and demographic spheres need intrinsic probabilistic models.

The difficulty of the statistical analyses of non-physical data, in fact, comes from what is the essential diversity of the relationships which must generally be highlighted. In practice, in contrast to physics, they always (or almost always) have to deal with relationships of functional dependence, because, even when the theoretical interpretation must rely on probabilistic schemes, the real (macroscopic) measures come from innumerable masses of homogeneous elements and lead to functional relationships. If, in order to produce homogeneous elements, the biologist or the demographer wants to divide the reference population according to all the characteristics which differentiate it, he would have to divide it into more and more restricted classes until it was reduced to single individuals. The probabilistic nature of the possible conclusions is therefore an aspect which relates not only to the expression of the results but also to the collection and elaboration of the data.

The physicist, whatever theory he creates, always knows it will have to be shown that future experiments confirm what his experiment has demonstrated. In other sciences which use statistics, observed facts which can be identically repeated do not exist, nor do they have, even when they are identical, an identical development.

Such considerations give rise naturally to the need to refer to more general mathematical models (of a probabilistic type) of which the schemes and models related to physical phenomena can be considered extreme cases. Fusionism is also opportune in the sphere of mathematics where the synergetic use of various (geometric, analytic, graphic) methods can be useful for the step by step solution of a certain problem.

Moreover, with such an approach it is possible to obviate another consequence of the present system: the development, sometimes even extreme, of analytical abilities to the detriment of synthetic abilities.

By proposing adequate courses, it is possible to teach something to individuals of any age; moreover, it is deleterious not to take the right moment to propose the best, most interesting *stimuli*. Children are good at using induction; it corresponds to the same mechanisms as linguistic (experimental) learning, which, as is well known, is much easier for younger people.

Moreover, by being accustomed to using inductive reasoning "early" they can avoid the hardening of the logical-deductive structures which are given priority in traditional teaching, even in scientific disciplines. Yet Italian schools show very strong trends toward conservation. The causes of this must be sought in the general situation in which the Italian school system operates, a situation which excludes competitivity, given that conservation is not assured by the quality of the product or by the efficiency of the system. Therefore any innovative proposal meets obstacles and involves high costs of various origins. It is impossible to leave these costs and the relative benefits of proposed innovations out of the analysis; hence the importance of a system of evaluation which can be used to monitor productivity and to suggest corrections useful for "adjusting the aim".

5. Variables of individual learning

Psychological investigators – as Fischbein (1975) reminds us – have found interest in the late development of the concept of probability, particularly in the growth-related development of children. What are the motivations for this delay, and for the recent (relative to the moment in which he was writing) interest? He lists some reasons, such as "the fact that behavioral phenomena are in themselves of an aleatory character", and that only recently has psychology arrived at the idea of needing probabilistic models to interpret behavior. Another is the "halo effect" produced by other sciences, contiguous to psychology, which are based on the theory of probability: in particular decision theory, information theory and cybernetics[1]. Since then, work and discussion on the axiomatic bases of probability have reopened the discussion on the meaning of the concept of probability itself, imposing a serious reconsideration of subjective features, or of the degree of confidence of every subject in events occuring within its field of interest.

Without doubt, Piaget and Inhelder's research[1] (see Piaget and Inhelder, 1951)) constitutes the first reference for this kind of study which subsequent authors have had to deal with. From our point of view, its didactic possibilities have not been completely taken advantage of in the current environment. Moreover they have not been directly elaborated by other authors, although they are always in the background, ready to guide those who want to consider them. For example, the result of the close relationship in the formation of the concept of probability as a function of age constitutes an undeniable guide for the organization of school work. It is also worth remembering the three stages of development identified by Piaget and Inhelder, because within age limits (which possibly can be shifted by usefully focused teaching) they represent a recognized background for the psychological genesis of the idea of chance.

In the first stage (before 7-8 years) a lack of operative abilities is observed, in the sense that the reasoning formulated by children at that age might be regulated only by a largely intuitive system, without set inclusion being managed, or even used. At this stage, the boy or girl might not distinguish between the possible and the necessary.

The second stage (from 7-8 to 11-12 years), characterized by the beginning of logical arithmetical operations, shows early embryonic developments of the idea of chance. Boys and girls at this age cannot manage to complete an exhaustive analysis of the combinatorial possibilities related to a certain situation. This is due to the fact that formal hypothetical-deductive thinking, which represents the substratum for formulating such a list of possibilities, is not yet completely developed. The children cannot support the techniques of reasoning which – according to Piaget – are the basis for the recognition (needed in the judgement of probability) that every possible outcome may occur.

---

[1] It could be interesting to remember the introduction of Piaget and Inhelder's book, in which we read: "Un mathématicien connu par ses travaux sur la théorie des probabilités" (he is probably referring to E. Borel) "nous souggérait un jour l'étude de ce problème psycologique: n'existerait-il en l'homme normal, une intuition de probabilité aussi fondamentale et d'usage aussi fréquent que, par exemple, l'intuition du nombre ?".

In the third stage (over 11-12 years), when combinatorial operations are understood and stabilized, a synthesis between combinatorial possibilities and probability can be given.

That is, children should be able to imagine the entire arc of sampled events and to represent to themselves that only some of those events, in subsequent experiments, may be realized. They are also able to distribute probabilities in this arc of events, by stating that some events will happen more easily than others.

Therefore, reviewing what has been one of the principal branches of the historical development of probability from 1500 (Cardano, Galilei, Pascal, Fermat,...), Piaget and Inhelder recognize a direct parallel between the achievement of combinatorial abilities and the development of the idea of probability in youth. In fact, the latter represents a prius of this: "(...) le jugement de probabilité s'organise en sa généralité *par un sorte de choc en retour* de l'opération sur le hasard" (see Piaget and Inhelder, 1951).

If Piaget's study only demonstrates combinatorial potential in teenagers, Fischbein (1975) explores the possibility that specific educational forms can accelerate the process of acquiring combinatorial abilities, and in the same way as those related to probability. "Probabilistic intuition" – he states – "is not spontaneously developed, except in a very restricted sphere. The understanding, interpretation, evaluation, and forecasting of probabilistic phenomena cannot be entrusted to simple uncultivated intuition, neglected and abandoned in a rudimentary stage of development, under the pressure of operative schemes which cannot articulate with them" (see Fishbein, 1975, p. 18).

Starting from Kahneman's and Tversky's analysis of the use of conceptual short-cuts (behavioral heuristics, according to the definition used in psychological literature), Agnoli (see Agnoli, 1991a,b; Agnoli and Krantz, 1989) also forcefully insists on the educational implications of these ideas. According to what she reports and has directly experimented with, it seems that the acquisition of logical concepts (substantially, as for Piaget, set inclusion and intersection, or implication and logical product in the language of propositions) appears stably in boys and girls from 11-13 years old and reduces the frequency of errors rather generally (see Agnoli and De Zuani, 1991a,b).

With regard to the origin of errors, this author also proposes to consider it, in adults as in children, as rooted in school teaching which too often does not recognize that certain decisions are taken in conditions of uncertainty[2]. If school teaching tends to emphasize the solution of problems of a deterministic character (that is, those problems for which just one correct answer is foreseen) and is not helpful in recognizing situations containing inherently uncertain results, then the children's intuition cannot open up to the presence of events to which a degree of personal confidence, a probability, can be assigned .

Another question, directly connected with what precedes, relates to eliminating the ambiguity of everyday language: if, for example, probable is confused with possible, then their respective negations, in the absence of semantic clarification, can lead to identifing improbable with impossible. But these unclarified linguistic features pervade large parts of mathematics teaching, as Baruk (1985) has shown with an abundance of particulars drawn from the 'homework' of her French students. She dedicates an entire section of her book to the existence of understanding and its relationships to the meaning which must be attached to a signifier.

6. Variables linked to school organization

School organization can be crucial to the possibility and incisiveness of the teaching of Probability and Statistics. In almost all educational systems such teaching, because of its recent introduction into academic programs, encounters all the problems related to every novelty (accentuated by the general ignorance of its basic concepts in almost the entire population) .

Often, one notes the lack of connection between different levels of study. For example, a daring reform regarding primary and middle school levels, which has promoted probability and statistics in the corresponding programs, has not been followed up in secondary schools, so that accumulated capital may be quickly lost, producing a rapid loss of acquired model-building abilities and a return to perceptual ones.

---

[2] A central problem of decision making in conditions of uncertainty, which are usually those which everyone must confront, is that individuals sometimes decide in ways very different from what economists and statisticians insist they should, that is, respecting prescriptions based on rational criteria. Notwithstanding this, one cannot affirm that individuals are irrational, indeed Descartes observed that intelligence is the good most widely spread among human beings.

Hungarian and also Italian experiences, for example, have shown and still show problems of the lack of vertical connection.

Moreover, a lack of horizontal connection is noted almost everywhere. A background reference analogous to that represented by historical synchronism, which is the supporting structure of teaching in the humanities complex, has not been widely used for the physical-natural disciplines.

In the last ten to twenty years some steps have been taken, above all for the linking of physical, chemical and natural sciences. Connections established between these disciplines and mathematics are less decisive, and practically non-existent with respect to Probability and Statistics, not withstanding that the latter could represent an important methodological adhesive.

On one hand, interdisciplinarity has had scarce attention in pre-university schooling; on the other, statisticians have invested little in the didactics of their discipline. It follows that in many cases the teaching of Probability and Statistics ends up being residual, unconnected with other disciplines, not even with mathematics where it is generally located. It is not by chance that the contents of Probability and Statistics are almost always introduced at the end of the academic year, or even of the educational cycle.

The construction of the curriculum, the definition of course prototypes, the planning of moments of interaction with other disciplines, the accumulation of a wealth of cases, are all goals to be pursued with great determination if we want to achieve results which are not banal or fragile.

Much national evidence emphasizes the importance of evaluation. The problem has different aspects. Firstly, it is clear that a system of evaluating and monitoring innovative experiences, such as that of teaching our subjects, is necessary to nurture the process of improvement.

We also need to consider that, in preparatory school, evaluation of the results of study is common. However, attention is directed at the final results of problems or exercises, and at the completeness of definitions. Instead, for our disciplines, and also because of their recent entry into teaching programs, priority should be given to evaluations relative to the process of reaching a result, and to the learning and use of models for reduction of problems. In other words, the evaluation of the learning of superior intellectual processes should be promoted, according to Bloom's taxonomo-cognitive objectives, which include the capacity to make applications, analyses and syntheses.

The actual teaching of Probability and Statistics is highly conditioned by the presence of corresponding components in the examinations. This is the evidence, for example, in the United Kingdom.

For our country, at the level of middle school, even though Probability and Statistics amount to 1/5 of all programs, examination problems on these topics comprise only in 5-10 % of the total questions. This certainly does not help.

One notes a divergence between modalities of teaching at university and pre-university levels in the experiences of many countries. Although pre-university teachers are generally trained in universities, when they return to schools they often, more or less quickly, regress to more common but not preferable schemes of teaching: from an interdisciplinary to a monodisciplinary approach; from an approach through models to perceptual evaluation; from an experimental attitude to a dogmatic one; from a predisposition to continuous up-dating to the defensive use of acquired training.

Until now, little has been studied about the presence of logical-deductive abilities and inductive abilities respectively in the school-age population. The experience of any teacher could testify that they are present in different proportions from student to student and this proportion is markedly linked to his/her original environment. Deductive abilities appear more widespread within higher social categories than within those less advantaged, as is immediately evident if one refers to the two categories of full-time students and part-time working students.

Generally, it is possible to argue that logical-deductive approaches to teaching can appear less democratic than inductive ones, to the degree that advantaged social classes are a minority in the school age population.

With regard to the teaching of Statistics and Probability, we need to add that it is largely justified by the contribution which it can make to understanding real experiences and problems. It leads to a preferring an inductive approach.

The link between the culture of origin and learning ability has been even less investigated.

To introduce a stimulating element for reflection, it is contended that more oppressive or dogmatic religions stimulate the acquisition of deductive abilities and hence predispose students to learning through axioms-theorems; while, naturalistic religions contribute to the expression of inductive abilities and predispose to learning starting from experience. On the other hand, if a culture or a religion forbids gambling, several possibilities for the learning of probability are precluded.

In the case of cultural minorities in which a non-standard language is predominant, such as in the ghettos of the United States, examples of a propensity to mathematical ovelinguistically based learning have been noted. Immigrants of very distant language families, such as Asian immigrants to the United States, have demonstrated great success in mathematical fields at university level.

These kinds of topics have been little considered in the debates in West European countries which have been relatively homogeneous from a cultural viewpoint. They will become more and more important however, as the flux of immigration from less developed to more developed countries increases.

7. Some suggestions

It would be useful to create a network for the exchange of experiences and information amongst centers, teachers and administrators. It is possible that many of the problems pointed out in the preceeding sections have been confronted locally, and that some results are already available. Knowledge of these could result in important and widespread effects.

ISI and IASE should promote research on the teaching of Probability and Statistics, taking into consideration the involvement of these subjects with other disciplines. Some possibilities have been indicated here. They need to be probed further, preferably in an international workshop, after analogous preparatory national workshops, of a very selective nature, with representatives of numerous disciplines. It could be an occasion to return to reflection on a wide horizon.

Bibliography

Agnoli F. (1991a), Teaching Implications of Misconceptions in Probability and Statistics (manuscript).

Agnoli F. (1991b), Development of judgmental heuristic and logical reasoning: training counteracts the representativeness heuristics, *Cognitive Development*, 6, pp. 195-217.

Agnoli F. and De Zuani E. (1991a), Fraintendimenti nel ragionamento probabilistico e implicazioni educative, *Età evolutiva*, XVIII, pp. 15-25.

Agnoli F. and De Zuani E. (1991b), Euristiche cognitive e ragionamento induttivo, *Giornale italiano di psicologia*, XVIII, pp. 193-215.

Agnoli F. and Krantz D.H.(1989), Suppressing natural heuristics by formal instruction: the case of conjunction fallacy, *Cognitive Psychology*, 21, pp. 515-550.

Baruk S. (1985), *L'âge du capitaine. De l'erreur en mathématique*, Seuil, Paris.

Fischbein E. (1975), *The Intuitive Sources of Probabilistic Thinking in Children*, D. Reidel Publishing Company, Dordrecht.

Piaget J. and Inhelder B. (1951), *La genèse de l'idée de hasard chez l'enfant*, Presse Universitaire de France, Paris.

THE CLASSIC DEFINITION OF PROBABILITY AND FRACTIONS: DIFFICULTIES AND POSSIBILITIES AT THE PRIMARY SCHOOL LEVEL

Carla Caredda
Dipartimento di Matematica, Università di Cagliari
Viale Merello 92, 09123 Cagliari, Italy
Maria Rosaria Puxeddu
Scuola Elementare Statale "Convitto Nazionale"
Loc. Terra Maini, 09100 Cagliari, Italy

1. Introduction

The concept of fraction represents, a moment of difficulty both for teachers, from whom it requires adequate teaching abilities, and for children, who must be highly motivated to understand it. It is not rash to state that even in the scholastic stages following the primary school level the approach to fractions requires special attention on the part of teachers and students. Fractional notation expresses different "mathematical moments" which obviously go beyond the traditional teaching concept of the *whole* to the *part*. The educative process at the primary school level traditionally starts from this approach and with this content, with the idea of latest broadening its fields of application. In some of these the same notation is subsequently used. In Italian primary school syllabuses for 1985, the goals outlined for the first cycle (first and second years) call explicitly for the use of the fraction symbol in the classic terms of the concept of *whole/part*:

> "... with the aid of an adequate number of objects, calculate the reciprocal connection double/half, triple/third, quadruple/fourth, and so on..."

Among the goals of the second cycle (*3rd*, *4th* and *5th* years) pupils must :

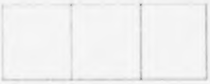
> "... find the fractions representing parts of suitable geometrical figures, sets of objects or numbers; conversely, given a fraction, find the corresponding part in appropriate geometric figures, sets of objects or numbers, paying special attention to decimal places. Compare and put in order the simplest fractions, using properly the number-line ...".

In the opinion of those who drew up the syllabus, this was the approach most suitable to the cognitive and psychological development of the child, who has already concluded his or her pre-operational period. It should form a solid foundation on which to build subsequent meanings of fraction symbolism. The extension of curricula to include probability, similitude and so on, does not appear to take for granted the priority given to the *whole/part* approach. In fact opens it to discussion and should stimulate research and experimentation to verify the validity of this assessment of priority. In school usage, the fraction symbol, in its meaning of whole/part, is found to be the most commonly accepted, and seems to establish a fairly natural relationship between child and concrete reality. To divide and recompose objects responds to the needs of children to act with and on things, and consolidates reversible thought. The mental capacity to arrive at a solution following determined pathways and then return, by means of analysis, to the starting point, represents the fundamental characteristic of reversible and operational thought. The development of an ever more complex society and the necessity of representing it symbolically obliges us to educate future generations in the decoding of different situations, even when they are expressed with the same symbols.

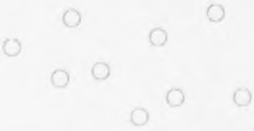
2. Fractions are difficult

That fractions are difficult is indisputable, and has been affirmed by others in the past (Hart, 1981). Historically, the traditional approach to teaching fractions has required the setting up of an itinerary calling for a more or less extended stage of manipulation and execution. Furthermore, psychological research into the learning process has shown that while it is true that perceptive experience helps in understanding this concept, it may also generate a certain functional fixity and form an obstacle to the enlarging of successive knowledge. The measurement of uncertainty is included, and quite rightly so, in a cognitive perspective that must be imparted to children at primary school level. How can we approach, gradually but systematically, the teaching of the analysis of probabilities? As is well known, this kind of analysis is expressed with fractional symbols signifying different things in different contexts. It is to be emphasized that the passage from action/perception to the fractional representation of numerical quantities has always represented a moment of crisis both for learner and teacher. Our experience has shown that after suitable drilling, children are capable of responding correctly to requests of the kind:

a) "Color two thirds of the figure"



b) "Draw a line around one third of these circles"



It goes without saying that the first exercise is much easier than the second, which implicitly requires the children to count the circles, divide them and single out the part to draw a line around. We should point out that this example also contains a perceptive difficulty, since the numerator of the fraction 1/3 does not correspond to the number of balls, which in this case should be three. This is quite different from the first example, in which children were asked to color two clearly visible parts. Still remaining within the limits of the concept of a "pie" (a rectangle or any other object that can be divided into equal parts) the fractional symbol maintains its undeniable difficulty.

3. Fractions and the classic definition of probability

At this point we ask: is it useful and effective to start from the whole/part approach in tracing didactic itineraries leading to the differentiated interpretation of the fraction symbol? In other words, is this the only possible way in which to reach an understanding of different realities? From the pedagogical viewpoint, it is not really right to use the fraction symbol if diversified didactic itineraries, one for each different situation, are not designed previously so as to ensure that the child comes up with a correct interpretation of reality. Experimentation with new teaching strategies can no longer be postponed if we wish to avoid gaps and cognitive lapses. The "new" use of the fraction symbol, with the inclusion of probability in the 1985 syllabuses, not only adds new difficulties to the traditional concept of fraction, but also represents a further obstacle between the learner and the concept of probability (Fischbein, 1969). When a child is asked what the probability is of an even number coming up on throwing a die, it is not a foregone conclusion that his or her answer will be 1/2. In fact, the representation in terms of a fraction of a probabilistic situation from the teaching viewpoint is not completely referable to the fractional representation of whole to part. At the end of their primary school careers, children should be able to understand and apply the classic definition of probability after graduated exercises (Boffa and Caredda, 1990). This means that they must be able to arrive at qualitative and quantitative evaluations of probability in simple, and obviously symmetrical, aleatory situations.

According to the classic conception, in order to express quantitative evaluations of probability one must be capable not only of performing combinatorial analysis, but must also know how to use the fraction symbol adequately. In this context, the fraction symbol is not an operator and does not express the division of the whole into parts, but assumes the function of measurer of uncertainty. If it is true that the whole/part concept is accessible to children, it is also true that it involves not only a certain difficulty in mastering it, but (Varga, 1969a,b) also a mental stereotype that may hinder future uses of this symbol.

Situations of uncertainty, for the very reason that they are uncertain, generate phenomena of different kinds. Firstly, we are faced with problems of a psychological nature. When we are required to make choices or come to a decision in situations in which we do not fully understand all their aspects and implications, our emotional states lie at the base of mostly "unreasoned" choices. Children who have not yet perfected their processes of maturation require both educative support that helps them to control their states of anxiety in situations of uncertainty as well as didactic support that accustoms them to making their evaluations as objectively as possible when exchanging and interchanging information. In the period in which children attend primary school they must learn to recognize the fact that in interpersonal relations the viewpoints of their peers have the same value as their own. It is in fact in this continuous confrontation, as well as in an adequate education, that it is possible to arrive at acceptance of the opinions of others and at the realization that others are of the same worth. Thus the capacity to concede merit to other viewpoints is the result of a long process that must not, however, be too long if we wish to arrive at a harmonious development of the personality. How closely affective-emotive development is connected with cognitive development is now an established fact in the literature of psychology, which stresses that acquiring a positive image of oneself is indispensable in overcoming failed attempts and negative results at an acceptable level of frustration. In school situations dealing with the themes of probability, it becomes necessary for children to be capable of controlling on the one hand the emotive state activated in aleatory situations and on the other hand of being able to express objective evaluations uninfluenced by personal desires and preferences. Concerning probability, the goals of the first and second years are:

> "In problematic situations taken from real life and play, use coherently and in the most meaningful way the expressions: 'perhaps', 'it's possible', 'it's certain', 'I don't know', 'that's impossible', and so on".

To trace a didactic itinerary that goes beyond the acquisition and use of the concepts of "true" and "false" to the concept of "possible" is certainly not an easy task. Teaching, to be effective, must bring ludic activities into play since they are congenial to the age of the child and its way of expressing itself and because their pedagogic value is universally recognized. This value is recognized in all scholastic disciplines and especially in mathematics, within which it not only facilitates the understanding of concepts, but also succeeds in verifying the theoretical assumptions referring to learning processes.

In aleatory situations, ludic activity maintains its pedagogic value, but it appears to us that it assumes a different role (Caredda and Puxeddu, 1990). While in common mathematical situations play verifies and completes theoretical assertions, in aleatory situations verification is lacking, but it gives rise to a cognitive conflict representing, in this context, the salient element in the learning process. The originality of the educative function of play lies in the "overcoming" of certainties and in the assumption of situations of uncertainty in cognitive structures. In ludic activities referring to the classic conception of probability (such as the throwing of a die) the outcome does not confirm expectations, thus often generating in the child the idea that the coincidence, or the lack of coincidence, between expectation and outcome is the result of good or bad "luck". The usefulness of the teacher's intervention consists of teaching that the rolling of "two" with the die is not the result of Johnny's prediction or of his "bad luck" if the two doesn't come up, but one of six equally possible outcomes.

Having attributed to play the role of stimulator of cognitive conflicts does not mean that we have eliminated other difficulties that are also present in situations of uncertainty. The pre-eminent function of language in teaching is universally recognized through its being present in all cultural contexts, to such an extent that it is rightly considered a "facilitator" in learning and teaching interactions. For it to be such, not only is semantic sharing a necessity, but also the two-way grasp of consciousness of what one intends to say. Acquisition in aleatory situations in terms such as: "perhaps", "it's possible", "it's certain", and so on must be placed within a learning process that must be indivisible linked to the

BEST COPY AVAILABLE

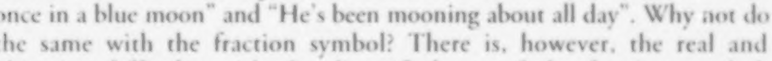
use of terms that are aleatory in daily experiences. The learning of terms that "speak of probability" must go beyond the scholastic and be inserted in a more general cognitive sphere. Teaching must move towards the overcoming of a linguistic separation between aleatory classroom situations and those of daily life and must, above all, see to it that the child accepts situations of uncertainty not only because they enrich his or her experiences (beyond the true-false dichotomy) but also to set him or her out in the direction of an evaluation of events. The use of the terms "perhaps", "it's possible", "it's certain" must be deprived of all emotive and affective connotations only to become more and more reasoned and objective in the course of a suitable didactic itinerary (Caredda and Puxeddu, 1991a).

4. Hypotheses concerning what to do: some considerations

In the present-day situation the use of the fraction symbol to express different meanings and the inclusion of the classic definition of probability in the whole/part fractional concept appear to us to be two possible approaches. Undoubtedly, the first approach requires the tracing of a gradual and continuous itinerary starting from the first year of primary school. The introduction of the generalized use of the fraction symbol in a teaching context would, in our opinion, lead to the elimination of those "certainties" that must in any case be overcome. At this point, we do not feel that a reference to Piaget (see Piaget and Inhelder, 1951) is out of place if we consider that it is in the period of concrete operations that the child succeeds in grasping and controlling at least two attributes of objects and situations. A didactic itinerary that fails to respect this rhythm of learning would lead to a cognitive development lacking, either wholly or in part, the logical tissue that is at the base of all human knowledge.

Therefore, the fractional symbol can be used effectively in "reading" different situations, as long as the child's learning processes are respected and graduated and continuous didactic itineraries are laid out. Such an idea may "frighten" primary school teachers, not only because, as we said before, fractions are difficult, but also because it will require them to widen their methodological and teaching horizons. But it should not be particularly difficult to get children to understand the efficacy and immediacy of a symbol, even when it is used to describe different situations. Today, the linguistic code in use is restricted, contrary to what Bernestein has stated (see Bernestein, 1970), as the result of sophisticated mental processes borrowed from computer science. We should not therefore be surprised if a single symbol is used to describe several different situations. The sharing of context, without which it is extremely difficult to understand others and to understand in general, is obviously indispensable. Linguistic experience shows that the same word in different contexts acquires different meanings. A primary school child, if he or she is aware of the different contexts, understands that the word "moon" assumes different meanings in the sentences "He wants the moon", "I win once in a blue moon" and "He's been mooning about all day". Why not do the same with the fraction symbol? There is, however, the real and objective difficulty with the diversified use of the fraction symbol, especially in consideration of the peculiarity of the learning processes the child goes through at the age at which he or she attends primary school. To this we must add difficulties connected with the periods of time needed for planning and developing the project itself. In the meantime, the continuity of educative interventions must be ensured. The didactic pathway must not be delayed or interrupted, and all useless and sterile superimpositions must be avoided by supplying a graduated and pedagogically planned procedure. In our view, the child of today, who is exposed to a large number and variety of signs and symbols and who knows how to play videogames should not find it difficult to recognize different contexts and interpret them with identical symbolic expressions.

If we asked children to represent the following cyclic cadence of shapes



and told them that such a cycle could be expressed not only with the numbers 2 and 3 (two representing the squares and 3 the circles) but also with the fractional symbol of two over three (2/3), it should not cause them any great difficulties. The same thing can be said for rhythmic sequences, which are often used in the first year in teaching reading and writing. The idea of blowing a whistle twice and beating a drum five times can be expressed both as (2, 5) and as 2/5. It should not be difficult, in situations describing spatial localization such as "Johnny went to see a play put on by his fellow fifth-graders. He sat in the second row in the third seat", to teach that this can also be expressed using the fraction symbol two over three (2/3).

The use of the fraction symbol in diverse contexts must not cause us to lose sight of the close connection between the symbol and the very concept of fraction, which represents the simplest expression of the subdivision of a whole into different parts. In elementary school syllabuses it is explicitly stated that geometric figures can be cut out or suitably colored, and that it is necessary to be able to recognize the complementary fraction, the proper fraction and the apparent fraction. The concept of fraction together with its graphic representation represents the most commonly known expression, although, as we have argued, the fractional symbol may express other concepts.

The second itinerary calls for the inclusion of the classic definition of probability in the whole/part acceptation of the fraction symbol. Despite the apparent difficulty this entails, we believe it represents a feasible itinerary, not least because it requires neither long-term planning nor the need for the conditions indispensable for the diversified use of the fraction symbol. Thus the introduction of the measurement of possibility also means to enlarge conceptually a process that is in any case already in step with the fraction itself. It is not so necessary to plan a special and different teaching approach as it is to familiarize children with the solving of problems having more than one solution. The classic problems with the drawing of the cartesian product can be a good beginning in orienting children towards finding as many solutions as they can. Not only this, but the use of the language of sets and the use of these in scholastic practice should help children to conceptualize "that the six sides of the die are a set and that the three even-numbered sides are one of its subsets". Introducing the measurement of the possibility of a certain event occurring fits into the fraction concept of whole/part without recourse to any teaching expedient. The idea is to propose and have the children verify that the six sides of the die are like a whole pie and that the three even-numbered sides represent, for example, three slices of pie that have been eaten. The symbolic representation is perfectly identical: 3/6 and thus 1/2. The setting up of teaching programs, and not only those concerning mathematics, must take into account appropriate situations that can aid in developing the capacity to compare and diversify knowledge. In the case of mathematics, this is generally favoured by combinatorial analysis, which systematically analyzes all possibilities. In our opinion, the operational proposals presented herein reflect the graduality and continuity necessary to ensure correct and meaningful assimilation (Boffa and Caredda, 1990).

Example 1: "During the spring track and field events, Al, Bert and Charlie got onto the podium. What positions could they have taken there?"

Example 2: "In a small lottery with only ninety tickets, the prize will be awarded to a ticket bearing a number from 1 to 90. Michael bought all the tickets numbered with multiples of five. Later, Dotty bought all the tickets numbered with multiples of three that were still left. How many tickets did Michael buy?

What are Michael's chances of winning the prize? Who has a better chance of winning, Michael or Dotty?"

Example 3: "A 40-card deck of playing cards has been divided into two stacks, one containing all the face cards and the other containing all the numbered cards. Both stacks have been carefully shuffled and placed on the table face down. From which of the two stacks are you more likely to pull out a heart? Explain your answer. What's the probability of picking the seven of diamonds from the stack of numbered cards? What's the probability of picking the king of clubs from the stack of face cards?"

Concerning the first example, we feel that all comment is superfluous, except to say that it can also be used in ludic and manipulatory activities and can be proposed even to children in the first two years of school. Up to the question "How many tickets did Michael buy?" the second example is an arithmetic problem. The need for analysis and determination of probabilities appears in the final two questions. In the third example, the most important aspect is represented by the fact that if the calculations are not performed it is impossible to discover that both choices lead to the same result.

5. Final observations

The ability to use the fraction symbol correctly in different contexts calls for a long-term teaching project which, through continuous feedback, must keep the teacher's work in tune with the child's developing psychological makeup. We agree with Bruner (1964) in saying that anything can be correctly and usefully translated into the way of thinking of a child of school age and that these first representations can be brought up later on and discussed more precisely and in greater detail. The idea of presenting the fraction as a single symbol with different meanings appears to us to be a feasible one as long as the conditions laid down are respected. The second approach involves fewer conditions, but it is no less important for this reason, since it is a question of including the analysis of probability, in the classic sense of the term, in the traditional concept of whole/part. In fact, the number of possible cases can be related to the "whole" while the number of most probable cases can be related to the number of parts taken into consideration. It is evident that the second approach is more feasible and is already in practice, though perhaps not always completely and knowledgeably, by teachers. The first approach is certainly less feasible owing to the long-term planning it entails, but it is also the one that carries with it a certain educative weight if we take into account its interdisciplinary and multidisciplinary implications. The possibility of transferring aleatory and non-aleatory situations from one cultural context to another and the possibility of using appropriately the different linguistic syntagmas are a part of the necessary and natural development of cognitive processes. In fact, independently of the hypotheses proposed herein, we feel it is important to underline the fact that the inclusion of probability in primary school syllabuses constitutes for children a new and further opportunity for learning and developing cognitive capacities and gaining the ability to judge, all acquisitions that are essential to life in a society of increasing complexity (Caredda and Puxeddu, 1991b).

Bibliography

Bernestein B. (1970), A sociolinguistic approach to socialization: with some reference to educability, in F. Williams (ed.), *Language and poverty: perspectives on a theme*, Markham, Chicago.

Boffa M. and Caredda C. (1990), *Probabilità e insegnamento elementare*, S.E.I., Torino.

Bruner J. (1964), *Dopo Dewey. Il processo di apprendimento nelle due culture*, Armando, Roma.

Caredda C. and Puxeddu M.R. (1990), *Il gioco: ostacolo o facilitazione nella comprensione di concetti probabilistici?*, *La Matematica e la sua didattica*, Armando, Roma, pp. 62-70.

Caredda C. and Puxeddu M.R.(1991a), *Ostacoli linguistici alla base delle dif-ficoltà nell'insegnamento-apprendimento della probabilità a livello ele-mentare*, Induzioni, n. 2, pp. 55-62.

Caredda C. and Puxeddu M.R.(1991b), *Una lettura critica su probabilità e statistica nei nuovi programmi della scuola elementare*, Induzioni, n. 3, pp. 87-92.

Fischbein E.(1969), *Initiation aux probabilités l'école élémentaire*, Educational studies in Mathematics, n. 2, pp.16-31

Hart K. (1981), *Children's understanding of mathematics: 11-16*, John Murray, London.

Piaget J. and Inhelder B. (1951), *La genèse de l'idée de hasard chez l'enfant*, P.U.F., Paris.

Varga T. (1969a), Combinatorials and probability for young children. Part I, *Journal of structural learning*, 1, n. 2/3, pp. 49-99.

Varga T. (1969b) Combinatorials and probability for young children. Part II, *Journal of structural learning*, 1, n. 4, pp. 139-161.

TEACH-STAT: A PROFESSIONAL DEVELOPMENT PROGRAM FOR ELEMENTARY TEACHERS (GRADES K-6) IN NORTH CAROLINA, U.S.A

Dargan Frierson
University of N.C. at Wilmington, Department of Mathematical Sciences
UNCW, Wilmington, N.C. 28403-3297, USA
Susan Friel
University of N.C., Mathematics & Science Education Network
201 Peabody Hall, UNC, Chapel Hill, N.C. 27599-3345, USA
Sarah Berenson
N.C. State University, 315 Poe Hall, Box 7801
NCSU, Raleigh, N.C. 27695-7801, USA
George Bright
University of N.C. at Greensboro, School of Education, Curry Bldg.
UNCG, Greensboro, N.C. 27412-5001, USA
Cliff Tremblay
Pembroke State University, Department of Mathematical Sciences
PO Box 2589, Pembroke, N.C. 28372, USA

1. Introduction

Teach-Stat is a professional development program for elementary teachers (grades K-6) in North Carolina funded by the National Science Foundation and jointly sponsored by the University of N.C. Mathematics and Science Education Network, the N.C. Department of Public Instruction and the N.C. Science and Mathematics Alliance. The Teach-Stat faculty are developing materials to help teachers expand their knowledge and understanding of statistics and to help them teach it to elementary students through an activity-based, data investigations approach. We expect teachers who participate in the Teach-Stat program will reframe the way they teach mathematics, science and social studies through the integration of data analysis activities into these subject areas. This paper provides an overview of Teach-Stat, a brief description of some of the workshop materials that have been developed for it, and a preliminary report on several research results that are beginning to emerge from the project.

2. Overview of Teach-Stat

The North Carolina *Standard Course of Study* (1989) for elementary school mathematics has for the past four years included a statistics "strand" in each of the grade levels 1 through 6, labelled as "data collection, display and interpretation" in grades 1 through 3 and as "graphing, probability and statistics" in grades 4 through 6. The competency goals in each grade expect students to be able to formulate questions, gather data using appropriate methods, analyze it in various ways (especially using graphical methods, but also with numerical measures in the upper grades), and interpret the results of their analyses both orally and in writing. Because most elementary teachers in N.C. have had minimal exposure to probability, statistics and data analysis during their preservice years in the University, little attention has been paid in many elementary classrooms to these important topics. To assist teachers in the development of their knowledge about statistics and how to teach it, the University of N.C. Mathematics and Science Education Network, through its Director Dr. Susan Friel, obtained funding for three years from the National Science Foundation for the development and implementation of the Teach-Stat project. This inservice, professional development program will not only help teachers expand their understanding of probability, statistics and data analysis, but will also help them learn appropriate ways to teach the concepts to elementary students, to reframe the way they teach mathematics in general (using a data analytic, "investigative" approach), and to incorporate data analysis and statistics into other subjects they teach, especially science and social studies.

During the 1991-92 academic year, University of North Carolina faculty representing nine of the sixteen constituent institutions of the UNC system met in several "seminar" sessions of 1-2 days each. These sessions served to orient the faculty to the content and research components of the project, to give them the opportunity to write the professional development curriculum they would teach from the following summer and to establish the evaluation and research agenda for the program.

During the spring of 1992, fifty-five elementary teachers were chosen from around the state of North Carolina to participate in a residential summer workshop. This workshop was held on a centrally-located college campus and for three weeks the participants were taught the activity-based curriculum in probability, statistics and data analysis which had been developed by the faculty during the previous months. The University faculty who taught the workshop were organized into four teams of 2-3 people and were assigned the responsibility of conducting the institute for two or three days at a time. Teaming permitted them to do joint planning with other faculty of diverse backgrounds (a statistician teamed with a mathematics educator and a science educator for example) and also to break the participants into smaller working groups. The elementary teachers were chosen to participate in this workshop based on recommendations of the Science and Mathematics Education Center directors at nine sites across the state, and the expectation was that they would be enthusiastic participants, very interested in learning about statistics and in teaching at least one unit on statistics in their classrooms during the coming academic year. This expectation was more than fulfilled by most of these "first-year" teachers as we visited their classrooms and conferred with them on numerous occasions during the 1992-93 school year. In fact, some teachers report that this project has "changed the way I teach". Not only do they appear to be incorporating more data analysis experiences throughout their teaching, but they also seem to be moving toward a more student-directed learning environment in general, in which they facilitate learning rather than direct it. This represents a change to a pedagogical style which more nearly approaches that described in the *National Council of Teachers of Mathematics Professional Standards for Mathematics Teaching* (1991).

During the 1992-93 academic year, the Teach-Stat faculty continued to meet for the purpose of revising the workshop materials while preparing for the second series of workshops to be conducted in the summer of 1993. During the second summer institutes, the first-year teachers were involved in teaching the workshop materials, along with the university faculty, to an additional 24 teachers at each of nine sites around the state. The involvement of the first-year participants in the second-year instruction was quite successful, especially since they could bring with them not only the knowledge gained in the 1992 workshop, but also the year's experience *doing* data analysis activities in their classrooms. They were able to provide much needed feedback to the second-year participants on specific investigations that were presented to children in their classrooms.

During Teach-Stat's third year, several more "seminars" will be held so the University faculty can write the final revision of the materials that have been used for the last two summer workshops. Two additional manuscripts will also be produced during the year and all are described in more detail in the next section. University faculty will continue to provide extensive and on-going support at each site for the first- and second-year Teach-Stat teachers and from each group of thirty will identify six who will be participating in the "statistics educators" component of the project. These statistics educators will then participate in an intensive one-week workshop taught by the University faculty during the summer of 1994 at their site to prepare them as regional staff development resource teachers in statistics/data analysis. There will also be a third offering of the summer workshop for an additional twenty-four elementary teachers at each of the regional sites, and the newly-trained statistics educators will work with the University faculty as part of the instructional team.

3. Materials produced by Teach-Stat

A minimum of three manuscripts will be produced by the Teach-Stat faculty and will form the core of the materials that could be used by others to replicate our program of professional development. In addition, other materials may be developed including video tapes of teachers participating in the activities during the summers or guiding their own students through data investigations in the classroom. No student materials are being produced, since most of the activities in the workshops are based on existing materials meant to be done by students in the classroom. Especially useful is the the series of *Used Numbers* books (see the Bibliography for specific titles).

The three manuscripts being developed are tentatively titled:

1. *Teach-Stat: Professional Development Program.* This manuscript will provide a "how to" discussion for planning and implementing a three-week teacher education institute. It will be written in a way that addresses teachers' needs in statistics education, but will be modelled after curriculum materials such as *Used Numbers* in terms of style and format. However, the audience for this manual definitely consists of those educators who are designing and teaching professional development programs for elementary teachers. The so-called "PCAI" format for the investigations in this manual is a modification of the four-stage model given by Graham (1987): Posing the question, Collecting the data, Analyzing the data and Interpreting the results. The model gives a framework to the reasoning process followed in most statistical problem solving and allows for the investigations to be structured in a consistent manner throughout.
2. *Teach-Stat: Investigations for the Classroom.* This manuscript provides a "how to" discussion for planning and implementing activities with elementary students for learning statistics, for learning mathematics using a data investigation approach, and for exploring applications in science and the social sciences using a data investigation approach. This is an "activities" document whose audience is elementary teachers. The activities will be designed to provide them with "investigation starters"

for a variety of purposes. Our first- and second-year teacher/participants in the summer workshops will be contributors to this document - if not actually writing materials, then certainly working closely with the University faculty to develop ideas which are then written by project staff.

3. *Teach-Stat: Statistics Educator Leadership Development Program.* This manuscript will provide a "how to" discussion for planning and implementing an intensive, one-week statistics educators' institute.

4. Preliminary research results

One of the major goals of the Teach-Stat research group has been to enlarge our understanding of the statistics knowledge and the pedagogical techniques for teaching statistics that elementary teachers bring with them to inservice programs such as ours. The obvious benefit of this increased understanding is that we will be better informed as we develop the materials and methods used in the Teach-Stat workshops. Very little attention has been paid in the literature to the statistics knowledge of teachers of any grade level (Shaughnessy, 1992), though the work of Mokros and Russell (1991) does begin to identify specific strategies that elementary teachers use to solve statistics problems involving application of the concept of mean. We know of no research on elementary teachers' pedagogical knowledge of statistics. Thus, our research focuses on developing a base of understanding about elementary teachers' conceptual and pedagogical knowledge of statistics. A brief description of some preliminary results of this research follows.

During the Spring 1992, the Teach-Stat University faculty developed, piloted and revised two instruments to measure elementary teachers conceptual knowledge of statistics and knowledge of statistics pedagogy. Three times in the last fifteen months, the original fifty-five first-year teachers (referred to as Core teachers) were administered the instruments: prior to the summer '92 workshop (referred to as the Pretest: Year 1); on the last day of that summer workshop (referred to as Posttest: Year 1); and at the end of the school year 1992-93 during which they integrated into their own classrooms the activities and knowledge they learned the previous summer (referred to as Post-Posttest: Year 1). The Post-Posttest results have not yet been fully analyzed, but will be used later to see if teaching statistics in the elementary classroom has had any additional effects beyond the summer workshops. In addition to the content and pedagogy surveys, concept maps were constructed by the Core teachers on the first day of the summer '92 workshop and in February 1993, and subsequently analyzed.

The statistics content surveys given to the Core teachers (both Pretest and Posttest) included twelve open-ended items which required them to read and interpret graphs (lineplot, histogram, boxplots and a scatterplot); answer questions about the concepts of mean, median, and mode; answer questions about what is "typical" about a given set of data and then make predictions from the data, among others. The items on the content survey were scored on a holistic basis using a scale of 0 (no response) to 4 (fully acceptable response). On the Pretest, there was a wide range of responses to each of the items (Bright *et al.*, 1993b). For questions related to graphing for example, only approximately one-third of the responses reflected good understanding of the content knowledge being assessed. Teachers' knowledge with regard to interpreting boxplots and scatterplots and with regard to measures of center was similarly weak with their justifications often incomplete and inaccurate. The Posttest saw some noticeable improvements in the clarity of explanations offered by the teachers, in part evidenced by greater use of statistics vocabulary. Comparison of the scores across administrations indicates a significant improvement on all but one of the items, which presented an ambiguous situation for the teachers' response. An especially interesting research effort arising from further analysis of the content survey data is focusing on efforts to determine if there are patterns to elementary teachers' development of ideas regarding the "center of the data" when presented with a graphical display of data (Berenson *et al.*, 1993c). Before instruction, teachers tended to fixate on "large" features of the graphs like the range of the data or the horizontal scale or the modes to explain "center" (the graphs consisted of a line plot in one item and a histogram in another). After instruction, teachers focused more on numerical measures obtained from the graphs, especially the median, instead of the graphical features themselves. Additionally, teachers' responses concerning "center" and "typical" were examined for stability over different graphical representations. Teachers were generally less stable in their fixations and ideas on the Pretest than the Posttest (Berenson *et al.*, 1993c).

On the statistics pedagogy surveys teachers were asked to discuss the statistics concepts that they thought were most appropriate for third graders and how they would teach them. In addition, they were asked to discuss their teaching strategies if students had access to manipulatives and/or calculators. Teachers' perceptions of the important statistics concepts shifted between the Pretest and the Posttest (Bright *et al.* 1993a) from what are classified as "isolated content" and "simple processes" (e.g., graphing, probability and organizing data) to "process" ideas (e.g., formulate questions, interpret data). Teachers seemed to have internalized the PCAI model used in the summer workshop, at least to the extent that their views changed about the statistics ideas appropriate for children to know. On the Pretest, most of the instructional strategies dealt with data gathering, graphing, and computing statistical values. On the Posttest a shift was noted in their teaching strategies to the use of a larger variety of data representations, extended discussion with children about the question formulation and follow-up questioning by the teacher to help children develop meaning for the interpretations of the data. This again reflects the instruction presented in the summer workshop, so the teachers seem to have internalized its goals. These shifts in perception are not surprising, however, since prior to the project the statistics background of most of the Teach-Stat teachers was weak and change was almost guaranteed. Further research is being conducted to measure additional changes in the statistics teaching strategies of our Core teachers after they have taught their own students during the recently completed school year and after they have taught their colleagues in the professional development workshops this summer.

The concept maps were examined to look at the ways elementary teachers represented their understanding of the broad area of statistics (Bright and Friel, 1993). The first analysis of these maps focused on categorizations of the statistics concepts included in their maps. Once the categories of concepts were determined, a map was coded as containing that concept if any of the key terms were found on the map. The percentage of maps which included each category of concepts was then computed. A great deal of diversity was found in their maps as evidenced by that fact that only three categories were included in more than 50% of the maps (collect data, measures of center, and tables/charts) and seven other categories appeared in 40-50% of the maps (counting, surveying, analyze data, graphs only, school graphs, normalizing summaries, and probability). In a second analysis of this data, special attention was paid to the teachers' understanding of the relationships among four critical statistics concepts in the N.C. *Standard Course of Study* (1989): graphing, collection of data, prediction, and measures of center. Only 3% of the maps included none of the four critical concepts, while 14% included one concept, 33% included two concepts, 36% included three concepts, and 14% included all four concepts. Graphing was included most often (91% of the maps) followed by collection of data (62% of the maps), measures of center (60% of the maps) and prediction (29% of the maps). As this analysis shows, graphing seems to be the only concept that is part of a large majority of elementary teachers' common understanding of statistics. It suggests that teachers may not be very familiar with, and may not be able to understand fully, the goals and objectives of emerging national and state curriculum frameworks. If these national and state standards are to be implemented, teachers must develop a better understanding of statistics concepts and we expect our research will help to identify areas of statistical thinking that need to be addressed explicitly in professional development workshops.

Bibliography

Berenson S.B., Friel S. and Bright G.W. (1993a), *Elementary teachers' conceptions of graphical representation of statistical data,* Paper presented at the annual meeting of the Research Council for Diagnostic and Prescriptive Mathematics, Melbourne, Florida.

Berenson S.B., Friel S. and Bright G.W. (1993b), *Elementary teachers' fixation on graphical features to interpret statistical data,* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, Georgia.

Berenson S.B., Friel S. and Bright G.W. (1993c), *The development of elementary teachers' statistical concepts in relation to graphical representations,* Paper submitted for presentation at the annual meeting of the Psychology for Mathematics Education - North American Chapter, Asilomar Conference Center, Pacific Grove, California.

Bright G.W., Berenson S.B. and Friel S. (1993a), *Teachers' knowledge of statistics pedagogy,* Paper presented at the annual meeting of the Research Council for Diagnostic and Prescriptive Mathematics, Melbourne, Florida.

Bright G.W., Berenson S.B. and Friel S. (1993b), *Statistics knowledge of elementary teachers,* Paper submitted for presentation at the annual meeting of the Psychology for Mathematics Education - North American Chapter, Asilomar Conference Center, Pacific Grove, California.

Bright G.W., Friel S. (1993), *Elementary teachers' representations of relationships among statistics concepts,* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, Georgia.

Graham A. (1987), *Statistical Investigations in the Secondary School,* Cambridge University Press, Cambridge.

Mokros J.R., Russell S.J. (1991), Toward an understanding of mean as "balance point", in R.G. Underhill (ed.), *Proceedings of the thirteenth annual meeting,* North American Chapter of the International Group for the Psychology of Mathematics Education, Vol. 2, pp. 189-195, Blacksburg, Virginia.

National Council of Teachers of Mathematics (1991), *Professional Standards for Teaching Mathematics*, Reston, Virginia.

North Carolina Department of Public Instruction (1989), Mathematics K through 8: Standard Course of Study, Raleigh, N.C.

Shaughnessy J.M. (1992), Research in probability and statistics: Reflections and directions, in D.A. Grouws (ed.), *Handbook of research on mathematics teaching and learning*, pp. 465-494, Macmillan, New York.

Used Numbers series:

Corwin R. and Friel S. (1990), *Statistics: Prediction and Sampling*, Dale Seymour Publications, Palo Alto, California.

Corwin R. and Russell S.J. (1990), *Measuring: From Paces to Feet*, Dale Seymour Publications, Palo Alto, California.

Friel S., Mokros J. and Russell S.J. (1992), *Statistics: Middles, Means and In-Betweens*, Dale Seymour Publications, Palo Alto, California.

Russell S.J. and Corwin R. (1990), *Sorting: Groups and Graphs*, Dale Seymour Publications, Palo Alto, California.

Russell S.J. and Stone Antonia (1990), *Counting: Ourselves and Our Families*, Dale Seymour Publications, Palo Alto, California.

ON THE PSYCHOPEDAGOGICAL ASPECTS OF TEACHING STATISTICS AT THE PRIMARY SCHOOL

Maria Pia Perelli D'Argenzio
Dipartimento di Matematica, Università Cattolica di Brescia
Via Trieste 17, 25121 Brescia, Italy

1. Introduction

The aim of this work is to present a few ideas on how to introduce the initial basic concepts of statistics; to date, studies on the teaching of basic concepts have generally been concerned more with the intrinsic aspects of the concepts themselves than with the problems involved in understanding them.

Before seeking to identify the problems involved in laying the foundation of basic concepts, we must bear in mind that all scientific research and knowledge starts out with a preliminary observation, eventually leading the scientist to formulate a hypothesis which is then followed by an examination of its validity. This well-defined course of investigation involves various stages in the mental process in which statistics can offer a valuable contribution. The aim of teaching statistics in schools is usually to provide students with tools useful in a variety of problem-solving situations. At the same time, using these tools also helps students to acquire a capacity for understanding, assimilating and interpreting information in general.

2. Teaching children the basic statistical concepts

The first step in understanding statistics is to identify the *population* being investigated and the corresponding *statistical units* it comprises. In primary school terms, this first step presents none of the difficulties of genuine research, because the type of investigation selected by teachers must be simple and must deal with easily identifiable, not too numerous, finite universes. These conditions are not always met, but a competent and sensitive teacher can easily satisfy such requirements.

The next step is to identify the *case* and the statistical *datum*. This datum has been clearly defined by the experts in terms of its intrinsic characteristics (continuous or discrete quantities, more or less easily distinguishable attributes, etc.), but the psychopedagogical problems it entails have been far less generally acknowledged. The process by which a datum is created has certain aspects which are not easily perceived by primary-school children: in particular, the type of abstraction views the statistical unit *not as a whole* but as an *anonymous element* lacking any feature or characteristic other than the one we are interested in. This type of abstraction is particularly difficult at primary-school age, which is a time for practical operations, not for formal reasoning; moreover, the child's intuitive faculties, which are essential to learning, typically take an all-inclusive rather than an analytical outlook. It is therefore important to help the child understand that, in the process of abstraction, the feature being dealt with is of interest in its own right, separate from all the other features which form a part of, and are combined together in, the statistical unit. For a child to understand that hair colour, age or weight, are separate statistical data which cannot and must not be identified with the individuals possessing those features calls for an abstraction which has more to do with logical processes than with practical considerations and is therefore not easy.

At the age considered in this study, it is also particularly difficult to understand that in observing certain features of each statistical unit, the intention is not to evaluate those features, but rather to draw conclusions about the presence of those features in the population as a whole, because statistical study aims to provide information on phenomena concerning the group, not the individuals.

To achieve this type of abstraction, it may be a good idea to plan a class activity similar to the following, which was used quite successfully in a third-year primary school class: each child brought a full-figure photograph of himself and was first asked to cut out the figure from the background, then to cut parts from the figure in order to separate certain features. The child sees his own image (which to his mind is representative of himself) robbed of the variety of features which combine to define it, so that it becomes an anonymous carrier of the feature and hence of the datum.

Another ability that the teacher has to establish in order for the child to consciously obtain the statistical datum is *to know how to see the alternatives* that exist in the *feature* being observed. When the feature is quantitative, and the alternatives can be expressed numerically, the task is fairly straightforward because the child has grown used to the comparison between numbers in learning arithmetic; but when the observed feature is an attribute, uniformity or diversity in the alternatives becomes a complex problem which a child is not always able to solve.

3. The problem of misinterpretation of statistical information

The second problem stems from the *tendency to re-attribute synthetic values to single statistical units.* To illustrate this point, I would like to quote a sonnet written by Trilussa, a poet in Rome at the turn of the century, which clearly expresses the layman's impression of statistical information and his inability to realise its informative value.

La statistica

Sai ched'è la statistica?
È 'na cosa che serve pe' fa' un conto in generale
del la gente che nasce, che sta male,
che more, che va in carcere e che spósa.

Ma pe' me la statistica curiosa
è dove c'entra la percentuale,
pe' via che, li, la media è sempre eguale
puro co' la persona bisognosa.

Me spiego: da li conti che se fanno
seconno le statistiche d'adesso
risulta che te tocca un pollo all'anno:
e, se nun entra ne le spese tue,
t'entra ne la statistica lo stesso
perché c'è un antro che ne magna due.

Roughly translated, the poem reads: you know what statistics is? It's something you use to take a general count of the people who were born, who get sick or die, who go to prison or get married; but what intrigues me is statistics that have to do with percentages, because then the mean remains the same for all, even for people who have nothing. Let me explain: from the way they count in statistics these days, it appears that everyone eats a chicken a year; and, even if you can't afford it, you come into the statistics anyway, because theres' somebody else who eats two.

Trilussa identifies as characteristic of statistics that they provide a means for describing social phenomena (people who were born, people who die, etc.); though this "definition" is obviously incomplete, it is nonetheless basically true. The poet then falls straight into the usual trap of attributing a synthetic value (the mean) to the individual who, together with the other statistical units, has contributed towards generating said value.

This misunderstanding is strongly rooted and is a further and almost fundamental justification for the introduction of statistical concepts at primary school level. To see the reason for this misunderstanding, we need to compare the learning of these statistical concepts with the learning of apparently similar mathematical concepts.

Learning a mathematical concept is a lengthy process which starts with observations that may appear complex, fragmented, and somewhat chaotic at first sight. This is followed by isolating a criterion on the basis of which several distinct objects can be considered "equal" among themselves but not equal to others. For istance, in introducing the geometrical notion of a rectangle, we overlook such features of an object as its colour, mass, hardness or material composition and consider only whether or not it has a particular form, which we call rectangular. We can further introduce the mathematical concept of the square by identifying a new set of four-sided objects that are at the same time equilateral and equiangular. This gives rise to a formula of equivalence which is what ensures that all the objects considered are interchangeable in terms of this criterion. Such equivalence among four-sided objects that are equilateral and equi-angular embodies the mathematical concept of the square in that these properties belong to all squares and only to squares. At this point, those properties that are characteristic of the square (in the mathematical sense) can be reattributed to all the real objects that led to the formulation of this concept.

In the formulation of statistical concepts, the first step again calls for a process of abstraction from the statistical units to statistical data and this first step is very similar to the process described above in the formulation of a geometrical concept. If, for example, we are interested in the shoe sizes of school children, we overlook the brand, type (sandals, shoes), colour or style of the shoe and consider size as our statistical datum for each child. The similarity stops here, however: when we go on to obtain decriptive statistics, which are synthetic values derived from the entire distribution of the data, we meet with conceptual difficulties. For one thing, while the arithmetic mean of a set of measurements has a precise physical interpretation, the same cannot be said of the arithmetic mean of a set of shoe sizes. Moreover, the synthetic quantity, e.g. the arithmetic mean, can no longer be re-attributed to each element of the population from which it was generated. In spite of that, young children (and adults too, as Trilussa's sonnet demonstrates) seem to have an instinctive tendency to follow their intuition and re-attribute these values to individuals. The child tends to personalise his experience, to refer all things to his own universe. "In medio stat virtus", says the Latin proverb, and the child considers this "middle" as exactly where he stands. If he detects any contradiction between the calculated "middle" and the position derived from his experience, he becomes convinced that the statistical processing of data produces a lie and is therefore of no use.

What is lacking here is an in-depth awareness of the fact that all statistical information, be it the synthesis or the totality of a number of individual items of information, *loses its individual connotation.* Attempting to use it in characterizing an individual is dangerous because it would involve a high degree of uncertainty. Proceeding with any inductive inference, i.e. drawing any conclusions from matters that are not directly observable, though their existence can be assumed, leads to the risk of error. Such inductive conclusions can only be interpreted correctly in probabilistic terms. Such people are confronted with decision-making situations of this type almost every day, it is essential for them to be trained to do so properly at the primary school level.

The primary school, then, is not required to train statisticians and provide high technical training; it should provide the qualitative education that would enable the child to overcome common, intuitive ideas. The final objective of a basic educational program could be, for example, to create a teaching unit capable of transmitting the two following concepts.

Statistical information is directly informative in terms of the community, not in terms of single individuals. It results from a combination of individual data which are generally variable, in that data from different cases are in general different. In addition, as soon as two diverse items of information are combined, the resulting synthetic information, by itself, cannot be reliably used to draw conclusions about the single individual observations.

But the tendency to re-attribute summary information back to individuals is common in adults and even more so in children, raising difficulties in understanding the meaning of statistical concepts, rather than in learning them. The risk inherent in teaching statistics in primary schools is that the child may tend to separate the notions learnt at school (surveys, calculation of mean values and other statistics) from what he really experiences and believes. In fact, when he discovers that what he has learned does not confirm his own experience (see Trilussa), he tends to rely with more conviction on past experience alone than on what he has "worked out" at school. He is unable to link subjective judgements and personal experience with what he is taught at school, which often leads to his accepting the theory at school (and being able to accurately calculate means, for example), but rejecting these skills for practical situations, and that means that the teaching method has failed. It is more important to implement teaching strategies designed to overcome these psychological difficulties than to provide technically valid instruction.

STAT-MAPS

Mike Perry
Department of Mathematical Sciences
Appalachian State University, Boone, NC 28608, USA

1. Introduction

STAT-MAPS, "Statistics-Materials and Activities for Problem Solving", is a four year project (1991-94) in the Department of Mathematical Sciences, Appalachian State University, USA. Funded by the National Science Foundation, the STAT-MAPS project is developing curriculum and materials for students in grades 9-12 (ages 15-17). The STAT-MAPS curriculum is giving attention to students with a broad range of abilities and interests, not just the college bound ones or the advanced students who have a special interest in science or mathematics.

The goals of STAT-MAPS are to: (1) describe a flexible curriculum for various secondary level settings, (2) develop effective instructional strategies for presenting this curriculum, and (3) provide materials for implementing the instructional strategies and curriculum. The project is based on the recommendations of The Curriculum Standards of the National Council of Teachers of Mathematics (1989) and builds on the previous work of the Quantitative Literacy Project (Scheaffer, 1986)

2. Foundations

General criteria

STAT-MAPS is developing materials that not only include topics, concepts and problems at levels appropriate for the secondary curriculum, but also include contexts for problems which are of interest to the secondary learner and relate to other areas of the curriculum. Instructional designs and strategies present statistics so that it is appropriable by the secondary student. Ideas are presented in a unified fashion so as to present a coherent picture of statistics.

Teaching and learning styles

Observers of instructional styles consistently report an increase in the use of lecture format and an accompanying decrease in small group or individual activity-based learning as the grade level increases. Formalization of procedures increases with grade level.

Learning theory research, however, suggests that the learning styles of most secondary students require exploratory activities with concrete realizations of ideas before abstract formalization of general concepts can be understood. STAT-MAPS is employing careful use of statistical problem-solving activities as a bridge across the concretion-abstraction gap suggested by current learning theory. STAT-MAPS activities encourage "intuitive" problem solving through a wide variety of real world problems. These concrete problems lend themselves to active, exploratory participation by the learner and may be used to nurture higher levels of reasoning.

3. The nature of the materials

The materials being developed by STAT-MAPS include:

1) Student text and activity manual;
2) Teacher manual.

DATA: to investigate our question about head sizes, we took measurements on 125 students (males and females) in five different classes. We measured head circumference (in millimeters) as described in the Eddie Bauer catalog.

ANALYSIS: the analysis uses stem-and-leaf and introduces the grouped frequency and relative frequency table and histograms.

INTERPRETATION: the pattern of head sizes and hence the shape of the histogram provide an answer to the question, and varying intervals is relevant to idea of a hat size.

COMMUNICATION: results are presented with tables, graphs, and written summary.

Example C

Problem type: Interpreting bivariate data
Levels: 2, 3
Title: Detectives, footprints and stature

Introduction: "Yes," he reported, after a short examination of the grass bed, "a number-twelve shoe, I should say. If he was all on the same scale as his foot he must certainly have been a giant". Sherlock Holmes, in The Adventure of Wisteria Lodge. Police officers often do not have accurate estimates for the height of a suspect in a criminal investigation. Footprints, either of the bare foot or more commonly as shoeprints, provide a possible way for estimating the height of a person. As suggested by inspector Sherlock Holmes, it is reasonable to expect that large persons have large feet. A lesser footprint would suggest a person of smaller stature.

QUESTION: is foot length a useful predictor of height?

DATA: to investigate the relationship between height and foot length, we took measurements on 121 students (male and female). Height and right foot length were measured with the subject standing barefooted.

ANALYSIS: visual analysis of scatterplots for trend is used. Fitting a line by the mean-mean method is introduced graphically and then the equation of the line is developed. Prediction error is described from the graph and the equation.

INTERPRETATION: the line is used for prediction, and prediction error is assessed.

COMMUNICATION: results are presented with tables, graphs, and written summary.

Example D

Problem type: Interpreting time series
Levels: 2, 3
Title: Murder rates in North Carolina

Introduction: there is great concern about the crimes rates in cities and regions of the United States. Many reports claim that the crime rate continues to rise. Most disturbing is the upward trend in violent crimes such as murder, rape, robbery, and assault. Many officials warn that we do not have enough policemen or jails to deal with the problems, and suggest that the criminal justice system needs additional support.

QUESTIONS: has the number of murders increased in North Carolina over the past ten years?

Can we predict what the number of murders will be next year?
DATA: uniform crime reports from the state of North Carolina Department of Justice provide data on several categories of crime, including the number of murders recorded monthly over the past ten years.

The teacher manual

STAT-MAPS is giving special attention to the teacher because statistics is a new subject in the U.S. curriculum. The limits of the teacher's knowledge about subject matter or about ways in which young people learn place predictable limits on what the teacher can offer and accomplish in the classroom. The teachers' beliefs about the nature of statistics are equally important influences on their students.

The STAT-MAPS Teacher Manual will include:

(1) an annotated version of the text to clarify objectives, provide supplementary discussion of statistical ideas and issues, and provide teaching ideas;

(2) a presentation and discussion of the various STAT-MAPS curriculum uses including class implementation schedules;

(3) a bibliography of teaching resources and references;

(4) suggestions on how to integrate other materials or resources such as videos into the STAT-MAPS curriculum.

Computing

STAT-MAPS is developing a curriculum for the future which employs computing resources. We are collaborating with Data Descriptions, Inc. to develop a special version of the DATA DESK software for school use.

4. The curriculum and instructional design

We agree with the statement of purpose for the textbook Introduction to the Practice of Statistics (Moore and McCabe, 1993): "We share the emerging consensus among statisticians that statistical education should focus on data and on statistical reasoning rather than on either the presentation of as many methods as possible or the mathematical theory of inference. Understanding statistical reasoning should be the most important objective". For secondary students, this may mean that sophisticated statistics are not used but what is used will be understood.

Curriculum levels

The STAT-MAPS curriculum is developed in four levels, roughly corresponding to grade levels 9-12. The materials are organized in a corresponding fashion. This organization provides flexibility so that they could be used in a curriculum which does statistics at each grade level, levels 1-2 could be combined within one year (grade 10), or levels 3-4 could be combined within one year (grade 11 or 12). The materials could be used in smaller pieces also.

Problem types / interpretation and inference

Ideas or activities are always presented through problem scenarios which may be classified according to several categories of interpretation and inference:

A. Fundamentals/Tables and Graphs
B. Interpreting Univarite Data
C. Interpreting Bivariate Data
D. Interpreting Designed Experiments
E. Interpreting Time Series
F. Estimating Population Values from Samples
G. Model Fitting
H. Statistical Decision Making

Statistics as a problem solving process

Statistics begins with the notion that we use data to answer questions. A statistics "problem" typically contains the four components:

(1) question
(2) data
(3) analysis, and
(4) interpretation and inference, in some order; and
(5) communication of results.

What makes a problem a statistical problem is the way we ask the question, the role and nature of the data, the peculiar ways that we look at the data, and the types of interpretations and inferences we make. STAT-MAPS uses this five point framework as a foundation for all of its text and activities.

Question formulation is an important part of any problem solving process. How we ask the question determines what kind of answer is attainable, and conversely, we must know what kind of answer is possible before we ask the question. The heart of the subject is data. Data is not just a set of numbers; data consists of a set of measurements. This distinction and the issues of measurement must be a major objective of statistical education. Understanding the role of data collection design must also be a major objective of statistical education. Statistical analysis, to a great extent, focuses attention on the variation of measurements and the pattern of variation in data is its typical starting point. Statistical education should develop an appreciation for this peculiar way that we look at data. The interpretation of the analysis and resulting inferences may lead to an answer for our question. Inference from data, whether formal or informal, is the final stage of the problem.

Understanding the nature and limitations of statistical inferences must be an important objective of statistical education. The communication of statistical findings in the form of graphical or tabular presentations as well as succinct written summaries should be emphasized. In each of the problem components, statistical reasoning must be employed. This "statistical reasoning" must be developed over time and the curriculum for high school students must continually cultivate and exercise this form of reasoning.

Student skills

STAT-MAPS categorizes student skills by concepts and techniques which are organized according to:

(1) Data Collection;
(2) Data Tabulation and Representation;
(3) Data Reduction, and
(4) Probability.

5. Example units

STAT-MAPS materials are organized by problem type and level. The presentation of ideas in the student text and the student activities are given in the framework of the problem solving process, with skills developed as needed for the given problem. This distinguishes the STAT-MAPS materials from most traditional textbook presentations which are organized essentially on the basis of techniques and methodology. By emphasizing statistics as problem solving, the development of statistical reasoning is quite natural because it is integral to the problem solving focus. The following are abstracts of STAT-MAPS units to illustrate how this is accomplished.

Example A

Problem type: Interpreting univariate data
Levels: 1, 2
Title: Reading Level

Introduction: magazines and newspapers are written for a wide variety of readers. The reading levels of publications vary. Some are written for persons who may have a limited education while others are designed for college graduates. The reading level may vary within a newspaper. An article on the sports page may not be at the same level as an article on the editorial page. There are several possible indicators of reading level that we can describe and analyze with statistics. Long words would indicate a high reading level, whereas use of short words would suggest a lower reading level. Other measures of reading level would include the lengths of sentences, the size of paragraphs, or the number of adjectives used.

We will compare the reading level of two newspaper articles by comparing the word lengths used in the articles. One article is by George Will, a political writer whose articles appear in different newspapers throughout the United States. An article about Whitney Houston is from The Star, a magazine which features stories on people in entertainment.

QUESTION: is there a difference between the word lengths used in the George Will article and the word lengths used in the article from The Star?

DATA: we examine and tally the word lengths for each of the three articles.

ANALYSIS: the analysis uses the ideas of max, min, median, mode, mean. It introduces frequency and relative frequency tables and bar charts as other ways of looking at the data which are relevant to the question.

INTERPRETATION: a comparison of word lengths is based on tables, bar charts, mean, median, and mode.

COMMUNICATION: the results are presented through the tables, graphs, and a written summary.

Example B

Problem type: Interpreting univariate data
Levels: 2
Title: Hat Sizes

Introduction: hats are made in a variety of styles and sizes. A merchant must decide what styles to keep in stock and how many hats of each size should be ordered. Are some hat sizes more common than others? If so, the merchant would want to order larger quantities of the more popular sizes. Hat size is apparently determined by head size. There are several possible measurements of the human head that might be used to describe head size. Mail order catalogs ask you to measure your head to determine your hat size. Here is the description from the Eddie Bauer catalogue (1992): "Hold tape level and firm. Measure circumference around the head, across temples and above eyebrow ridges".

QUESTION: are some head sizes more common than others?

ANALYSIS: visual analysis of time series plots and prediction from linear trends are used. Prediction based on non-linear trends is introduced.

INTERPRETATION: the notion that this type of prediction depends on assumptions about future trends is emphasized. A line may be fitted to a portion of the series, and the portion used affects the prediction.

COMMUNICATION: results are presented with tables, graphs, and a written summary.

6. Project status

Class Testing of materials was conducted at two sites during the spring of 1993. Two classes at Garinger High School in Charlotte, an urban environment, used STAT-MAPS units. One class in the rural community of Yanceyville also participated in testing. These will be expanded to four sites during 1994. The design and development of the DATA DESK software just began during the summer of 1993. It is anticipated that the project work will be completed during the spring of 1995.

The principal persons in the STAT-MAPS project are: Mike Perry, Project Director; Gary Kader of Appalachian State University, Associate Project Director; Peter Holmes of the University of Sheffield, England, consultant.

Bibliography

National Council of Teachers of Mathematics (1989), *Curriculum and Evaluation Standards for School Mathematics*, Reston VA.
Moore D.S. and McCabe G.P. (1993), *Introduction to the Practice of Statistics*, W.H. Freeman and Company, New York.
Scheaffer R. (1986), The Quantitative Literacy Project, *Teaching Statistics*, 8(2).

STATISTICS: A MOTIVATOR FOR SECONDARY SCHOOL MATHEMATICS

Richard L. Scheaffer
Department of Statistics, University of Florida
Gainesville, Florida 32611, USA

1. Background

The notion that data collection and analysis should be an integral part of the secondary mathematics curriculum in the United States has enjoyed an ongoing debate since at least the early part of this century. It was suggested by some around the time that Walter Shewhart introduced the control chart for quality improvement in the 1920's and was still being suggested by many in the years following World War II. At that time, one of those proposing that data analysis skills are essential requirements for all students was W. Edwards Deming who, of course, went on to become famous for his work in process improvement. This practical side of mathematics education lost momentum in the 1960's with the advent of the so-called "new math" movement, which emphasized formal mathematical structure and logic.

Not giving up easily, those who thought that students should be exposed to real, everyday applications of mathematics continued to press for more statistical concepts and techniques to be included in the school mathematics curriculum. One of the first organized efforts in this regard came about in the late 1960's with the formation of the Joint Committee on the Curriculum in Statistics and Probability by the American Statistical Association (ASA) and the National Council of Teachers of Mathematics (NCTM). Through publications of materials for a general audience, such as high school students and teachers, and workshops for teachers, this Committee kept honing the conceptual setting and the pedagogical style that would make statistics an acceptable, and even welcome, addition to the mathematics curriculum. Statistics as an educational strand, kept alive by the work of the Joint Committee and others through the 1970's, blossomed into new and vibrant life in the 1980's. This new life for statistics in education parallels the strong interest in process improvement techniques developing in industry during this time, and this duality between the educational and industrial roles for statistics seems to be more than coincidence.

A synergistic mixture of educational policy recommendations, new emphases in statistical practice, and the computer laid new groundwork for statistics in K-12 education over the last decade. With regard to policy, the major shift in ideas on mathematics education is, perhaps, best summarized in the report *Everybody Counts*, issued by the National Research Council in 1989. This report points out that a transition must be made from preoccupation with routine skills to "developing broad-based mathematical power", as in "using tables, graphs, spreadsheets, and statistical techniques to organize, interpret and present numerical information". Simultaneously, NCTM issued its *Curriculum and Evaluation Standards for School Mathematics*, which contain a carefully delineated strand in statistics throughout the curriculum, as well as an emphasis upon modeling from data in the study of algebra and functions. The data analysis tone of the Standards can be seen in the following quote.

> In this age of information and technology, an ever-increasing need exists to understand how information is processed and translated into usable knowledge. Because of society's expanding use of data for prediction and decision making, it is important that students develop an understanding of the concepts and processes used in data analysis.
>
> This standard should not be viewed as advocating, or even prescribing, a statistics course; rather, it describes topics that should be integrated with other mathematics topics and disciplines.

These recommendations for adding statistics to the curriculum are best implemented by having students actively involved with data collection and analysis, and this approach links directly to the emphasis upon data exploration in the practice of modern statistics. The tools of exploratory data analysis are largely graphical, thus requiring the use of a computer for efficient implementation. For the first time, educational policy, the practice of statistics by professionals, and the computer revolution seem to be working in harmony with one another. Data analysis appears to be far more than a passing fad in mathematics education.

Moving to more specific observations, it is true that those writing the statistics strand for the NCTM *Standards* developed many of their ideas around the work being done at that time by the ASA-NCTM Joint Committee. This work was mainly through the Quantitative Literacy Project (QLP), a project funded by the National Science Foundation (NSF) to develop curriculum materials and provide workshops for teachers in the area of statistics for the secondary mathematics curriculum. The Data-Driven Curriculum Project (DDC), which is the focus of the remainder of this report, can be thought of as an extension of the QLP in the directions recommended by the *Standards*.

2. Motivation and need

The technological society of today is dependent upon the collection and interpretation of data for virtually all of its decision making processes. Sample surveys measure consumer preferences and the unemployment rate. Clinical trials determine which new drugs are allowed on the market. Statistical process control improves manufacturing processes and helps keep an industry viable in a competitive world economy. Everyone, then, must understand something about data collection and interpretation if they are to make informed decisions on economics, jobs, politics, health and even entertainment. It is vitally important that the students of today be taught data analysis skills before they become the workers and leaders of tomorrow.

In addition to being important in their own right, data analysis skills help the students build connections between mathematics and other subjects in the school curriculum and to the world outside of the classroom. Data on genetics can lead to interest in the biological sciences. Data on the motion of the stars and planets can lead to interest in the physical sciences. Data on demographics and voting patterns can enhance interest in the social sciences. Data on industrial productivity can strengthen interest in business and economics. These connections can change student attitudes about mathematics and its usefulness as a tool in practical problem solving. Mathematics then becomes both relevant and interesting. Along the way, data collection and summarization can hone basic mathematical skills while data interpretation and model building are developing higher order thinking skills.

Historically, integrating data analysis into the mathematics curriculum has not been a common practice and has not been part of the educational background of mathematics teachers. The materials and inservice workshops of the QLP and similar projects that have sprung up over the past ten years have been well received by teachers, many of whom have attempted to teach the material to their students. Virtually all of these teachers testify to the fact that it takes great effort on their part to fit statistical ideas into the taught curriculum, and some never do succeed. Clearly, there is a need for curriculum materials that embed the important statistical ideas into the traditional mathematics topics so that students see the benefit of both.

3. Goals and objectives

Motivated by the success of the QLP and mindful of the continuing need for curriculum materials and teacher enhancement in the area of statistics, ASA and NCTM joined forces once more to plan and carry out the Data-Driven Curriculum Project (DDC). The QLP provides materials in certain areas of data exploration, probability, and inference for proportions arising from sample surveys in a style that makes the material accessible to teachers and students at the secondary level. It is not, however, comprehensive in its coverage of statistical topics and it does not allow for easy implementation into the traditional mathematics curriculum. The DDC expands the coverage of statistical topics so that a more comprehensive view of the field emerges and introduces the statistical ideas as motivation for the mathematical topics to which they are tied most closely. Most of the mathematical topics emphasized in the Standards (mainly algebra, functions and geometry) can be illustrated and motivated in this way. Through this two-pronged approach, teachers will be able to use data analysis to meet two objectives: teach a comprehensive body of knowledge in elementary statistics and motivate the major topics in mathematics through real applications.

The specific goals of the DDC are to:

a) Deliver a detailed scope and sequence plan for a data analysis strand in the high school (9-12) mathematics curriculum that implements the NCTM Standards by developing practical connections to the main concepts of algebra, geometry and functions.

b) Develop a series of curricular modules fitting the scope and sequence plan in a manner that ensures the coverage of a comprehensive body of knowledge in statistics as well as motivation for many topics of the mathematics curriculum.

c) Identify materials, including books, computer software, videotapes and manipulatives, that can be used to support the data analysis strand.

d) Produce a series of workshops for state mathematics supervisors, district coordinators and lead teachers for the purpose of disseminating the results of this project in a manner that will lead to implementation in the schools.

To date, a series of nine modules has been outlined and working drafts have been written for seven of these. Field tests of three modules have been completed and the next four will be field tested in the fall of 1993. A series of five workshops will begin in the fall of 1993. All work is to be completed by the end of 1994.

4. Overview of materials

A brief overview of each of the modules under development, along with the current working titles, is provided below. Much of the work is presented in the form of activities for students working in cooperative groups, with the instructor acting as a catalyst and coach, whereby the students can "discover" key mathematical and statistical ideas on their own.

Our numerical world introduces students to the notion of looking at numbers in the context of what they represent (data) and always thinking about how to interpret these numbers. The unit is built around developing the concepts of size, unit, variability and bias. Activities range from collecting and interpreting data on roller skating injuries to deciding which compact disk player to purchase.

Symbolic expression is built on the fact that a key component of mathematical reasoning is the ability to think abstractly about quantitative concepts so that ideas can be generalized beyond the numbers at hand. This abstraction depends upon symbols and their use in expressions, equations, and functions. A "spreadsheet" formulation of problems is used so that symbols acquire a natural meaning as column labels. Symbolic expressions then show how one column is related to another, independent of the data points in any particular row. Examples include work with wage and hour log sheets for employees in a small firm, discussions involving rates commonly found in vital statistics (birth rates, death rates, accident rates), and development of standard statistical measures of center and spread.

Linear relations builds on the notion that many relationships between continuous variables are studied in reference to a linear function. Data that exhibit approximate linear association are used to illustrate slope as a rate of change. Two such relationships lead to the problem of solving systems of linear equations. Applications relate to changes in automobile prices over time, the association between heights of children and heights of their parents, the relationship between points in a basketball game and percentage of shots made, and the relationship between fat and caloric content of fast foods, to name a few.

Exploring equations extends the concepts begun in the preceding module to develop the idea of an optimal (or "best") line to express the linear relationship between two variables, based on the data. Quadratic functions and their properties are introduced in the context of the least squares fitting procedure, where optimal properties are illustrated without reference to calculus. Correlation is also discussed.

Mathematical modeling with functions brings together the study of various kinds of functions and the idea of building a mathematical model suggested by data. Students are encouraged to make conjectures about models and then verify or refute those conjectures. Geometric relationships, direct and inverse variation, exponential and logarithmic functions, and proportionality are the mathematical topics of emphasis. Fitting models to data, transforming data, and studying patterns in residuals are some of the statistical ideas considered. Examples include exponential growth and decay of certain populations, relationships between weight and size of animals, and growth of plants over time.

Centers considers the notion of "average" from a geometric perspective, noting that most definitions of "center" actually minimize some distance measure. Weighted averages and the centroids of planar figures are given meaning through realistic applications such as determing a basketball players performance rating and locating a warehouse to minimize travel distance to stores.

Frequency tables and probability shows that data in the form of relative frequencies for categorical responses can be interpreted as approximate probabilities for predicting future events of a similar nature. Thus, marginal, joint and conditional relative frequencies are used to motivate similar concepts for probability. Testing hypotheses (chi-square test) is introduced here through simulation arguments, and simulation is used to introduce a variety of probability distributions for discrete random variables. A number of examples are drawn form clinical trials experiments and drug screening tests.

Advanced modeling develops matrix algebra through extending many of the modeling notions from earlier modules to more complex cases. Statistical models that go beyond simple linear regression are presented, as are probabilistic models involving Markov chains.

Projects outlines the essential steps that cooperative groups should consider when attempting a project that goes from posing a question, through data collection and analysis, and on to a written interpretation of results. Guidelines are given for designing sample surveys and experiments, with the latter concentrating on "clinical trials" and factorial arrangements of treatments. Students are encouraged to critique articles from the media that include statistical arguments.

Taken together, the nine modules cover statistical topics on exploring data, collecting data through planned studies, and inference. These are spread out over modules that could be used from grade 9 through grade 12 in covering a majority of the mathematical topics deemed worthy of enhanced emphasis by NCTM. It is hoped that this project will successfully demonstrate that statistics and mathematics can be integrated into a modern curriculum and will serve as a model for future curriculum development.

5. Summary

The development of a data-driven curriculum for high school mathematics appears to be in line with the needs of students to see more motivation and application within the mathematics classroom and to develop important skills to carry beyond the classroom. The revised curriculum under development is designed to raise the quantitative literacy of all students as it builds connections among mathematics, science, and technology. It models a new approach to the teaching of mathematics, the approach required by the NCTM *Standards*, as it emphasizes hands-on activities for students and discovery of concepts through data. Technology in the form of graphing calculators and computers is an integral part of the teaching and learning style being promoted through these materials and workshops. This project attempts to connect topics of importance in a modern mathematics curriculum to a modern view of statistical science for the purpose of enhancing student interest and skills in both areas.

TEACHING PROBABILITY AND STATISTICS AT UNIVERSITY LEVEL

Invited papers	George Box Lennart Råde
Contributed papers	Carol Joyce Blumberg Megan Clark John D. McKenzie, Jr. Hristo Nikolov *et al.* Jenö Reiczigel Karl Stielau Roel van Strik Kamen Velev *et al.*

WHAT ENGINEERS NEED TO LEARN ABOUT STATISTICS

George Box
Center for Quality and Productivity Improvement
University of Wisconsin-Madison
610 Walnut Street, Madison, Wisconsin 53705, USA

1. Introduction

I will begin by asking three questions and supplying three brief answers which are elaborated on in the rest of my talk. The questions are:

a) Why
b) What — should we teach engineers, scientists (and statisticians)
c) How — about statistics?

I believe the answers are:

a) To catalyze and robustify
b) Methods which catalyze and robustify — the process of problem solving and scientific discovery
c) Engage the student in

2. Why?

Ensuring that graduating engineers and scientists are familiar with statistical methods of design and analysis is, in the United States at least, an uphill battle (see, for example, Bisgaard, 1991). Remarkably little headway has been made in *requiring* that these techniques are taught to students in engineering and the sciences. This is partly because many scientists and engineers regard the statistics that they *have been* taught as totally irrelevant to the problems they face. I believe that they are often quite right about this.

2.1. *Discovery is an iterative process*

Scientific investigation has two aspects: *discovery* (problem solving) and *testing* the solution. Consider a scientific investigation intended to provide a drug which can cure a particular disease. Such an investigation involves: a) the *discovery* of an effective and manufacturable chemical substance; b) the *testing* of this substance to ensure its effectiveness and safety for human use.

The process of discovery must be undertaken in the same spirit as a detective solves a mystery and finds a criminal. It is necessarily a *sequential iterative* procedure.

Testing the final product is a much more formal process. It parallels the trial of the accused within very strict rules of admissible evidence. It is usually a *one shot* affair.

Unfortunately, the modern statistician, who is often also the teacher of engineers and scientists, has frequently been trained only for the role of designer and analyzer of the one shot trials appropriate to testing the solution after the work of discovery has been done. Consequently, at least in recent times, statistics has often not been allowed to play its critical role as a catalyst to the process of discovery itself.

Emphasis (Tukey, 1977) on the importance of exploratory data *analysis* addresses this problem. However, exploratory *inquiry* involves not only data analysis but the whole process of investigation and, in particular, the *sequential use of designs.*

2.2. *Continuous never ending improvement*

We can understand the critical importance of sequential investigation if we consider a central principle of modern quality technology - that of "Continuous Never Ending Improvement". This idea seems at first to be in conflict with the law of diminishing returns. Suppose, for example, you have a response curve like Fig. 1(a) or more generally a response surface defined by $y = f(\mathbf{x})$ and you want to find the levels of $\mathbf{x}$ which maximize y. When you have (nearly) achieved this maximization, shouldn't experimentation stop?

This reasoning applies to a fixed model. But (Box, 1993) in real investigations neither the functional form of the model, nor the identity of the variables $\mathbf{x}$, nor even the nature of the response y is fixed. They *evolve* as new knowledge comes to light. As is illustrated in Fig. 1(b), experimenters must be allowed to learn as they go. Whereas a fixed model leads inevitably to the barrier posed by the law of diminishing returns, the developing model provides for expanding returns and the possibility of *never ending improvement.*

In practice, some level of improvement will be adequate for present purposes and the improvement process will temporarily halt; but this quiescent period will last only until external circumstances again renew the incentive for change.

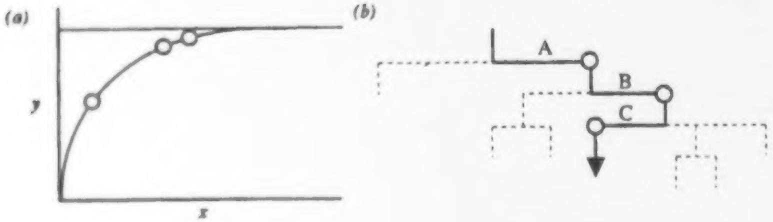


Figure 1. (*a*) *Diminishing returns from the fixed model* $y = f(x)$. (*b*) *Potentially expanding returns from the evolving models* $\mathbf{y}_A = f_A(\mathbf{x}_A) \rightarrow \mathbf{y}_B = f_B(\mathbf{x}_B) \rightarrow \mathbf{y}_C = f_C(\mathbf{x}_C)$. *Dotted lines are "roads not taken".*

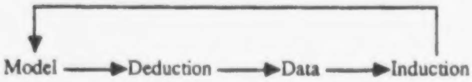
3. What?

The data needed to solve a problem may, at different stages of investigation: a) *already exist* - for example in the library, or in past operating records; b) be obtainable by *observing* but not interfering with the operation of the process; c) need to be *generated* by experimental intervention.

3.1. *Data from passive observation*

We can regard any operating system (whether it be a system for admission to a hospital or a system for producing transistors) as continually generating potentially useful information - rather as a wireless transmitter transmits radio signals. But just as radio receivers are needed to hear radio signals so tools of analysis are needed to understand what the process has to tell us about how it can be improved. The simpler (graphical) techniques such as those employed in Ishikawa's (1976) seven tools are of particular importance. This is because not only can the engineer use them, but s/he can also teach the whole workforce how to use them. The idea is illustrated in Fig. 2 which shows the "seven tools" plus a few more.

Whether by such simple methods or by more sophisticated techniques, effective problem solving necessarily requires a sequential approach. In such a strategy each step uses information gained at previous steps to follow an iterative course.


Model
Deduction
Data
Induction

This process of learning as you go involves a constant interaction between statistical considerations and subject matter (engineering) know-how. As the investigation proceeds it frequently turns out that the data suggests ideas (new variables, different responses, new levels of variables, etc.) that were *not* in mind the beginning of the investigation.

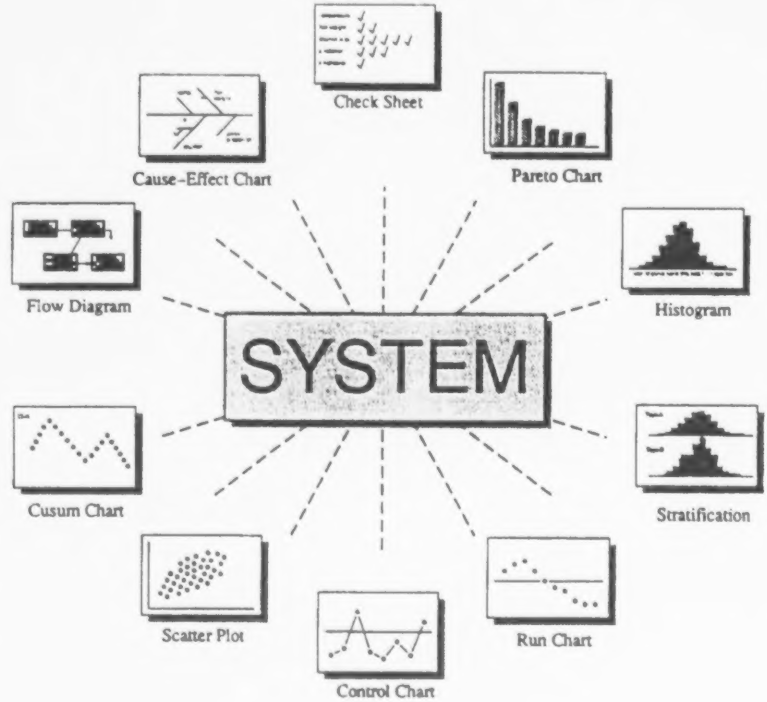

Check Sheet
Cause-Effect Chart
Pareto Chart
Flow Diagram
Histogram
SYSTEM
Cusum Chart
Stratification
Scatter Plot
Run Chart
Control Chart

Figure 2. *Some simple statistical tools for quality improvement by the workforce. From top left: Flow Diagram, Cause-Effect Chart, Check Sheet, Pareto Chart, Histogram, Stratification, Run Chart, Control Chart, Scatter Plot, Cusum Chart.*

Statisticians who have been trained only to run one shot trials may be uncomfortable with the idea that they should teach engineers that scientific investigation requires an indeterminate and flexible model, the responsibility for whose evolution must be shared with the experimenter. Certainly in the United States, the role of scientific iteration in statistics is usually dealt with by pretending it doesn't exist and discussing only one-shot investigations. By doing this statistics can be rigorously mathematicized. Unfortunately the process of scientific induction which cannot be modeled mathematically is the only way in which truly *new* ideas can be introduced. By cutting a living process of investigation in two, you kill it.

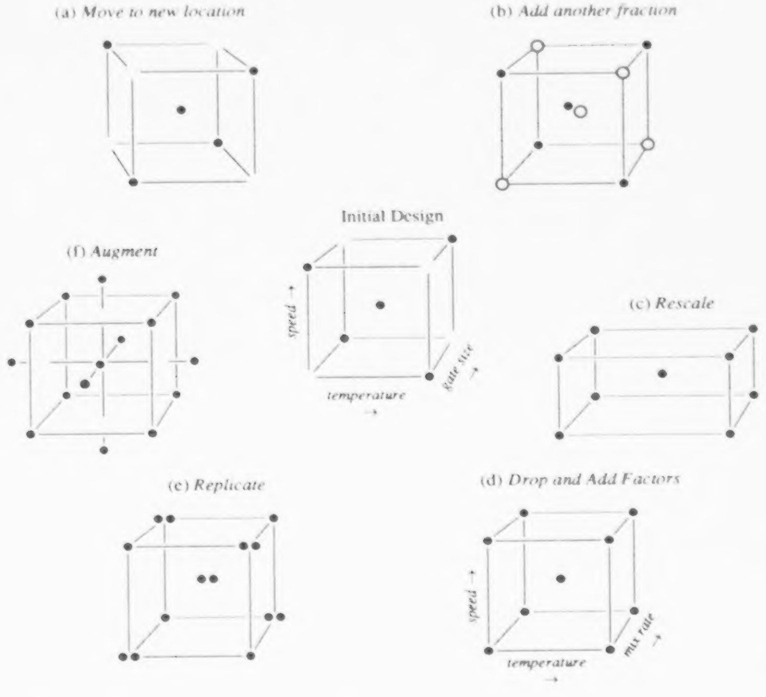

(a) Move to new location
(b) Add another fraction
Initial Design
(f) Augment
speed →
gate size →
temperature →
(c) Rescale
(e) Replicate
(d) Drop and Add Factors
speed →
mix rate →
temperature →

Figure 3. *Illustration of sequential experimentation for the three variable case. Depending on the results from the initial design and subsequent designs, various different courses might be taken.*

3.2. *The use of sequential experimentation*

For problems that cannot be solved by using data already available or obtainable from passive observation of the process designed experimentation is needed. At the beginning of an investigation the experimenter knows least about the *identity* of the important variables, the *location* of the experimental region of interest, the appropriate *scaling* and *transformation* of the variables, the degree of sophistication required to *model* the system and so forth. Thus "one-shot" experimentation which attempts to cover all bases with a single large experiment planned at the beginning of the investigation is likely to be extremely inefficient. Sequential experimentation with associated sequential assembly of designs which at each stage builds on information already obtained is usually much better.

3.3. *Informed extrapolation*

It must be made clear to the student engineer that the scientific process, involves not only the investigational iteration shown on the left of Fig. 4 but also *informed extrapolation* indicated on the right of that figure. Such extrapolation can be from the small scale to the full scale, from one location to another, from one investigator to another, and so on. As has been emphasized by Deming (1950, 1986), except in enumerative studies the final link with practice is not made using statistics or formal probability. It is made by using technical judgment. It is important to understand however that the *basis* for this extrapolative technical judgment can be very strong or very weak depending on how the investigation was conducted. Consequently, although no absolute guarantees are possible, by taking certain precautions in the design process we can make the job of informed extrapolation less perilous. The precautions for strengthening the analytical and judgmental links that connect investigation and application

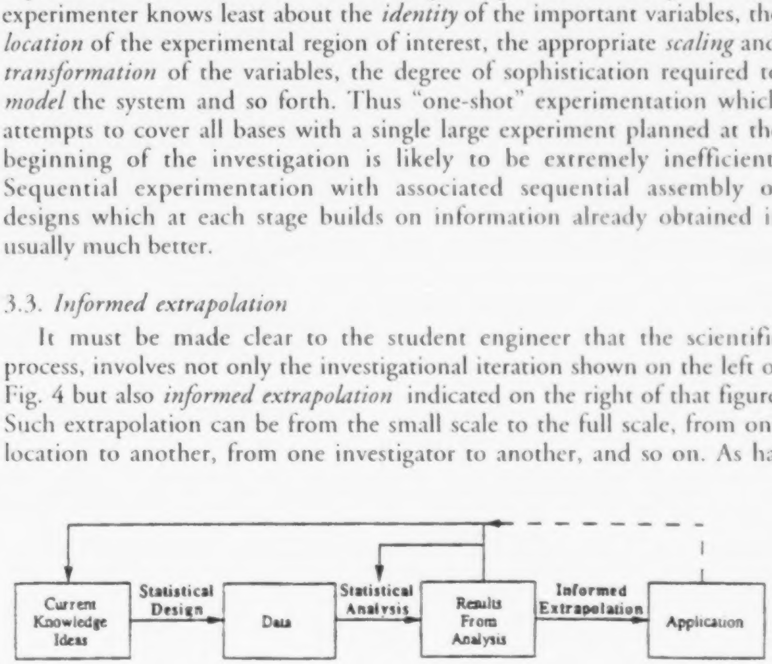


Figure 4. *Feedback and linkages in the process of problem solving and in application of the results.*

involve issues of *philosophy, analysis,* and *design* and have been discussed more fully elsewhere (Box, 1993). Briefly we need to consider:

Philosophy	a) The investigational process as an iteration
Analysis	b) The state of control and null behavior of the system
	c) The contribution of "distribution free" analysis
	d) The need for exploratory data analysis
Design	e) The purposes of planned experimentation
	f) The role of randomization and blocking
	g) The advantages of comparative experimentation
	h) The rationale of factorial and orthogonal design
	i) The purposes of robust design

(*a*) *Nature of the investigational process.* Iterative investigation, of the kind illustrated in Fig. 1(b) and Fig. 4, itself makes an important contribution to extrapolability.

(*b*) *State of control and null behavior of a system.*The IID state is one that does not occur naturally but can sometimes be approximately induced artificially, for example, by a steady elimination of the larger disturbing factors and/or by feedback control (see, for example, Box and Kramer, 1992). Whether any real process has ever existed in the IID state seems dubious. While every effort must made to ensure that an experimental system is in the best state of control we can get, we would be sanguine indeed to rely absolutely on an assumed state of control to make our experimental conclusions valid.

(*c*) *"Distribution free" analysis.* The role of deductive mathematical analysis in providing a wider base for extrapolation has been greatly overrated. In particular, the so called "distribution free" tests, although free of the assumption of normality, continue to make the same devastatingly dangerous (see, for example, Box and Newbold, 1971) *distributional* assumption that errors are independent or at least interchangeable.

(*d*) *Exploratory data analysis.* Such analysis is of great value at every phase of an investigation, both for providing clues to unexpected phenomena and for warning of situations where standard analysis will be misleading.

(*e*) *Need for planned experimentation.* Fisher, encouraged by Gosset, realized in the 1920's that it would be necessary to *design* experiments to address specific questions of interest. Two concepts were of particular importance - minimizing experimental error and maximizing extrapolability. Fisher's (1935) principles of planned experimentation are designed to solve these problems simultaneously.

(*f*) *Randomization and blocking.* Randomization makes the statistical conclusions robust against any kind of non IID disturbance and in particular against a non-stationary disturbance (see, for example, Box, 1990).

(*g*) *Comparative experimentation.* Experiments which most strain extrapolation are those concerned with an *absolute* measurement. Comparative experiments, from which you wish to learn if A and B are different and by how much, are less troublesome. Whenever possible, therefore, a problem of absolute measurement should be transformed into one of relative measurement. For example, in the manufacture of certain automobile parts, a robot grasps each item as it comes off the line and makes a series of measurements. However, these measurements are not compared with standard values. Instead, after measuring each manufactured item, the robot moves back to measure a standard "perfect part" which is available for continuous reference. The *differences* in the measurements of the manufactured part and the standard part are used to decide whether or not the part is in conformance. The advantage of this procedure is that you do not have to have the robot in a perfect state of calibration - any bias will be equally reflected in both the measured part and the reference part.

(*h*) *Factorial and orthogonal experimentation.* Factorial experiments are simultaneously statistically efficient and provide estimates of interactions; in addition they can be run so that the advantages of randomization, blocking and comparative experimentation are maintained. But Fisher (1935) also had in mind the questions of extrapolation and robustness. He remarked "(extraneous factors) may be incorporated in (factorial) experiments designed primarily to test other points; with the real advantages that, if either general effects or interactions are detected, that will be so much knowledge gained at no expense to the other objects of the experiment; and that, in any case, there will be no reason for rejecting the experimental results on the ground that the test was made in conditions differing in one or other of these respects from those in which it is proposed to apply the results".

These ideas were extended further by Youden (1961a,b) who was then working at the Bureau of Standards to develop what he called "rugged" methods of chemical analysis. His experiments used *fractional* factorial designs (introduced by Finney, 1945) to further increase the number of extraneous factors that could be tested.

(*i*) *Robust design.* Thus the concept of using statistics to design a product that will operate well in the conditions of the real world clearly has a long history going back at least to Gosset in the early part of the century. Early *industrial* examples are due to Morrison (1957) and to Michaels (1964). We owe to Taguchi (1986) demonstration of the wide industrial importance of these robust design ideas.

In the courses we teach engineering students at Madison, we discuss all the above points which impinge on extrapolation but, in teaching the section on robust design, we do not employ Taguchi's techniques (Box, Bisgaard and Fung, 1988). Instead we use what we believe are simpler and more efficient methods which, in particular take account of the important points made by Morrison and Michaels. Also, in short courses, we have successfully taught these ideas to engineers and scientists in many parts of the world.

4. How?

Of prime importance to the engineer investigator is the philosophy of the sequential generation of appropriate data. In particular different design approaches are required at different stages of investigation:

Screening designs. When we do not know which of the number of possible factors are the important ones fractional factorial designs and other *orthogonal arrays* (Plackett and Burman, 1946) are of great importance for screening out what Juran has called the *vital few* factors from the *trivial* many.

Empirical modeling. Sometimes, possibly as a result of previous screening or because of previous knowledge, the important variables are believed to be already known but we wish to find out how they affect a particular response or a number of responses. At this stage of experimentation, factorial designs, mild fractions and *response surface methods* (see, for example, Box and Wilson, 1951 and subsequent publications) are particularly important. The models used at this stage are more or less empirical and are often based on polynomials sometimes in suitably transformed variables.

Mathematical modeling. When sufficient physical knowledge of the system is available *mechanistic model building* techniques may be used. Possibly as a consequence of previous empirical experiment, the process functions derived from a supposed physical mechanism are employed. Such relationships are frequently represented by differential equations or integral equations. They often employ numerical methods of solution and non-linear least squares and non-linear design, together with model checking and model discrimination techniques (see, for example, Box and Draper, 1987; Bates and Watts, 1988).

Experimental design and analysis can also have different *objectives*. Among these are:

i) to raise (change) the mean value of some quality characteristic
ii) to reduce the variance
iii) to find conditions which in some sense produce a *robust* product or process.

4.1. *Hands-on experience*

In teaching the art of catalyzing scientific investigation with statistical methods the most vital component is *hands-on* experience. One way of achieving this (Hunter, 1977) is to require individual projects in which students carry through investigations using statistical methods at home and in the lab. In addition some simple experimental device can be employed for demonstration in the classroom. We have found the paper helicopter shown in Fig. 5(a) to be a very convenient means of teaching experimental design and analysis at many different stages of instruction (see, for example, Box, 1992).

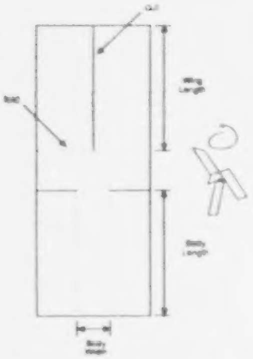
In this first example I will show how a "play acting" scenario may be used to teach the class some fundamental ideas. Three student volunteers are needed to play the parts. I will call these Tom, Dick and Mary.

Tom stands on a ladder and drops the helicopter from a height of twelve feet or so while Dick times its fall with a stopwatch. We explain to the class that we would like to find an improved helicopter design that has a longer flight time.

Mean and variation. We start by Tom dropping a helicopter made from blue paper. He drops it four times. We put the results up on the overhead projector and we see that the flight times vary somewhat. This leads to a discussion of variation and to the introduction of the average as a measure of central tendency and the range and standard deviation as measures of spread.

Comparing mean flight times. At this point Dick says "I don't think much of the blue helicopter design. I made this red helicopter yesterday of a different design. I dropped it four times and got an average flight time considerably longer than you just got". So we put up the two sets of data - the four runs made with the blue helicopter and the four runs made with the red helicopter and show the two averages and standard deviations. Then we demonstrate a simple test that shows that there is indeed a statistically significant difference in means, in favor of the runs made with the red helicopter.

Validity of the experiment. Then Mary says "So the difference is statistically significant. So what? It doesn't necessarily mean it's because of the different helicopter design. The runs with the red helicopter were a)



b)

FACTORS		–	+	*EFFECT*
Paper Type	(*P*)	Regular	Bond	0.13
Wing Length	(*W*)	3.00"	4.75"	0.77
Body Length	(*L*)	3.00"	4.75"	-0.40
Body Width	(*B*)	1.25"	2.00"	-0.02
Paper Clip	(*C*)	No	Yes	0.05
Fold	(*F*)	No	Yes	-0.10
Taped Body	(*T*)	No	Yes	-0.15
Taped Wing	(*M*)	No	Yes	0.17

Random Order	Standard Order	*P*	*W*	*L*	*B*	*C*	*F*	*T*	*M*	Flight Time
7	1	–	–	–	–	–	–	–	–	2.5
13	2	+	–	–	–	+	–	+	+	2.9
4	3	–	+	–	–	+	+	–	+	3.5
9	4	+	+	–	–	–	+	+	–	2.7
1	5	–	–	+	–	+	+	+	–	2.0
12	6	+	–	+	–	–	+	–	+	2.1
15	7	–	+	+	–	–	–	+	+	2.9
3	8	+	+	+	–	+	–	–	–	3.0
6	9	–	–	–	+	–	+	+	+	2.4
16	10	+	–	–	+	+	+	–	–	2.6
14	11	–	+	–	+	+	–	+	–	3.2
5	12	+	+	–	+	–	–	–	+	3.7
11	13	–	–	+	+	+	–	–	+	1.9
10	14	+	–	+	+	–	–	+	–	2.2
2	15	–	+	+	+	–	+	–	–	3.0
8	16	+	+	+	+	+	+	+	+	3.0

c)

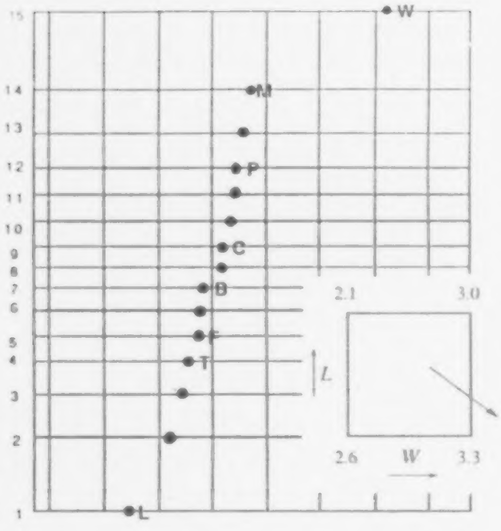


Figure 5.

a) *Construction of a paper helicopter.*

b) *Results from 16 run fractional factorial experiments showing the factor levels and the calculated main effects of the eight factors.*

c) *A normal plot of the effects from the helicopter experiment. The inset diagram summarizes the conclusions.*

BEST COPY AVAILABLE

made yesterday when it was cold and wet, the runs with the blue helicopter were made today when it's warm and dry. Perhaps it's the temperature or the humidity that made the difference. What about the paper? Was it the same kind of paper used to make the red helicopter as was used to make the blue one? Also, the blue helicopter was dropped by Tom and the red one by Dick. Perhaps they don't drop them the same way. And *where* did Dick drop his helicopter? I bet it was in the conference room, and I've noticed that in that particular room there is a draft which tends to make them fall towards the door. That could increase the flight time. Anyway, are you sure they dropped them from the same height?". So we ask the class if they think these criticisms have merit and they mostly agree that they have, and they add a few more criticisms of their own.

We tell the class how such considerations led Fisher some seventy years ago to consider the precautions necessary in running an experiment so that it can provide data which leads to unambiguous conclusions rather than to an argument. We show how his ideas can be used to compare the blue and the red helicopter by making a series of paired comparisons. Each pair (*block*) of experiments involves the dropping of the blue and the red helicopter by the same person at the same location. The decision as to which helicopter should be dropped first is made *randomly* by tossing a penny. The conclusions are based on the *differences* in flight time within the pairs of runs made under identical conditions. We go on to explain however that different people and different locations could be used from *pair to pair* and how, if this were done it would strengthen the extrapolability or as Fisher (1935) said "widen the inductive basis" of the experimental results. If the red helicopter design appeared to be better, one would, for example, like to be able to say that it seemed to be better no matter who dropped it or where it was dropped.

A fractional factorial design. As another example of the use of this device; at a later stage in the course it is used to illustrate the value of fractional factorial designs as screening devices. It is supposed that a brainstorming session by an engineering design team has led to the selection of eight factors to be studied. These selected factors are listed at the top of Fig. 5(b) together with the two conditions (indicated by minus and plus signs) at which each will be tested. It is thought likely that only a few of these factors will have important large effects. We are thus in the familiar "Pareto" situation where, as Dr. Juran says, we want to screen out "the vital few from the trivial many". The design used in Fig. 5(b) is a 2^{8-4}_{IV} fractional factorial. The student may be taught something of the theory of these designs; however, to use them, all they really need is a table such as has been supplied by Bisgaard (1988) which gives this and other eight and sixteen-run designs with a succinct description of their properties and analysis. As is well known, the 2^{8-4}_{IV} design has two very valuable characteristics:

a) if there are interactions between pairs of factors they will not bias any of the 8 main effects of the factors;

b) if only up to 3 factors are of importance, the design will produce a complete 2^3 factorial design replicated twice in those three important factors no matter which ones they are (see, for example, Box, Hunter and Hunter, 1978).

Flight times for the sixteen helicopter types obtained from an experiment run in random order are also shown in Fig. 5(b). From these flight times, 8 main effects and 7 strings of two-factor interaction may be calculated on the assumption that interactions between 3 or more factors may be ignored. The effects are plotted on probability paper in Fig. 5(c) suggesting that real effects are associated with W (wing length) and, less certainly, L (body length). On the basis that the remaining effects falling around the straight line are mostly due to noise, we can summarize the data simply in terms of the inset diagram in Fig. 5(c). The experiment immediately provides not only an improved helicopter design but also indicates the direction in which further experimentation should be carried thus initiating a sequential process of experimentation which can be carried as far as one desires.

Another aspect of this approach is highlighted by discussing with the class whether they are satisfied with *flight time* as the sole criterion. In earlier lectures we have emphasized to the class that what happens in each run of an experiment must be carefully documented - for example the fact that helicopter #7 hit the table leg and that the run had to be repeated. The need for careful observation is emphasized, perhaps leading to the conclusion that an additional criterion such as flight *stability* should be considered in future experimentation. This teaches the lesson that appropriate and feasible objectives *cannot* always be determined in advance. The *criteria* to be used in assessing the results may need to be modified or totally changed during an investigation as more is learned. The helicopter can also be used to illustrate the process of iterative experimentation at later stages. For example (a) to demonstrate the use of experimental design to reduce variance, (b) to illustrate the importance of estimating variance components and of using the appropriate error term and (c) to consider the problems of obtaining a robust product or process.

Acknowledgment

This work was supported by a grant from the Alfred P. Sloan Foundation.

Bibliography

Bates D.M. and Watts D.G. (1988), *Nonlinear Regression Analysis and Its Application*, John Wiley & Sons, New York.

Bisgaard S. (1988), *A Pratical Aid for Experimenters*, Starlight Press, Madison, Wisconsin, USA.

Bisgaard S. (1991), Teaching Statistics to Engineers, *The American Statistician*, 45(4), pp. 274-282.

Box G. (1990), Must We Randomize Our Experiment, *Quality Engineering*, 2(4), pp. 497-502.

Box G. (1992), Teaching Engineering Experimental Design with a Paper Helicopter, Quality Engineering, 4(3), pp. 453-459.

Box G. (1993), Statistics and Quality Improvement, *Journal of the Royal Statistical Society*, Series A, 156, pp. 209-229.

Box G., Bisgaard S. and Fung C. (1988), An Explanation and Critique of Taguchi's Contributions to Quality Engineering, *Quality and Reliability Engineering International*, 4, pp. 123-131.

Box G. and Draper N. (1987), *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York.

Box G.E.P., Hunter W.G. and Hunter J.S. (1978), *Statistics for Experimenters*, John Wiley & Sons, New York.

Box G. and Kramer T. (1992), Statistical Process Monitoring and Feedback Adjustment-A Discussion, *Technometrics*, 34(3), pp. 251-285.

Box G. and Newbold P. (1971), Some Comments on a Paper by Coen, Gomme and Kendall, *Journal of the Royal Statistical Society*, Series A, 134, pp. 229-240.

Box G.E.P. and Wilson P. (1951), On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society*, Series B, 23, pp. 1-45.

Deming W.E. (1950), *Some Theory of Sampling*, John Wiley & Sons, New York.

Deming W.E. (1986), *Out of the Crisis*, MIT Press, Cambridge, MA.

Finney D J. (1945), Fractional Replication of Factorial Arrangements, *Annals of Eugenics*, 12, pp. 291-301.

Fisher R.A. (1935), *The Design of Experiments*, Oliver and Boyd, Edinburgh and London.

Hunter W.G. (1977), Some Ideas About Teaching Design of Experiments, with 2^5 Examples of Experiments Conducted by Students, *The American Statistician*, 31(1), pp. 12-16.

Ishikawa K. (1976), *Guide to Quality Control*, Asian Productivity Organization, Tokyo. Available in USA from UNIPUB, New York.

Michaels S.E. (1964), The Usefulness of Experimental Design (with discussion), *Applied Statistics*, 13(3), pp. 221-235.

Morrison S.J. (1957), The Study of Variability in Engineering Design, *Applied Statistics*, 6(2), pp. 133-138.

Plackett R.L. and Burman J.P. (1946), The Design of Optimum Multifactorial Experiments, *Biometrika*, 33, pp. 305-325.

Taguchi G. (1986), *Introduction to Quality Engineering*, White Plains, New York.

Tukey J.W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading, MA.

Youden W.J. (1961a), Experimental Design and ASTM Committee, *Materials Research and Standards*, 1, pp. 862-867. Reprinted in H. H. Ku (ed.), *Precision Measurement and Calibration*, Vol. 1, Special Publication 300, Gaithersburg, MD, National Bureau of Standards, 1969.

Youden W. J. (1961b), Physical Measurement and Experimental Design, *Colloques Internationaux de Centre national de la Recherche Scientifique* No. 110, le Plan d'Expériences, pp. 15-128. Reprinted in H. H. Ku (ed.), *Precision Measurement and Calibration*, Vol. 1, Special Publication 300, Gaithersburg, MD, National Bureau of Standards, 1969.

SIMULATION AND THE STUDENT ENGINEER

Lennart Råde
Chalmers University of Technology
S 412 96, Gothenburg, Sweden

1. Introduction

The aim of this paper is to stress the importance of the simulation of random situations in today's engineering world and that for this reason simulation should be treated in all basic stochastics courses and especially so in such courses for engineering students. R. W. Hamming mentions in Hamming (1993) that in the 1950s at Bell Lab one out of ten experiments was made on the computer and nine out of ten in the laboratories but that later the situation was reversed with nine out of ten experiments made on the computer, that is, by simulation.

For some engineering students, for example, students of computer science engineering or industrial engineering and management, it might be appropriate to have a specific course on simulation.

Simulation can also be used as a didactical tool to give the students concrete experience and background for abstract concepts and specific methods in stochastics.

Simulation is, according to Ralston and Reilly (1983), "the representation of certain features of the behaviour of a physical or abstract system by the behaviour of another system".

The following is, according to Law and Kelton (1991), a list of some particular problems for which simulation has been found to be a useful and powerful tool.

1) Designing and analyzing manufacturing systems.
2) Evaluating hardware and software requirements for a computer system.
3) Evaluating a new military weapons system or tactic.
4) Determining ordering policies for an inventory system.
5) Designing communication systems and message protocols for them.
6) Designing and operating transportation facilities such as freeways, airports, subways, or ports.
7) Evaluating designs for service organizations such as hospitals, post offices, or fast-food restaurants.

8) Analyzing financial and economics systems.

Some of the reasons why simulation is used in these and other areas of applications are the following.

1) An analytical solution is not possible.
2) An analytical solution is too complicated.
3) Simulation can be used to estimate unknown probabilities, distributions, expectations and so on.
4) Simulation can be used to give a concrete illustration of what random variations and stability can be expected in a given system.
5) Simulation can be used to check an analytical solution.
6) Simulation can be used as a didactical tool.

Note that even if it is possible to find an analytical solution to a problem, for example in queueing theory, it might still be worth while to simulate the situation to give a concrete picture of what random variations or stability can be expected in the system.

Monte Carlo methods are closely related to simulation. By Monte Carlo methods are usually meant methods of solving deterministic mathematical problems by experiments with random numbers. For example, complicated multiple integrals can be estimated by simulation.

Simulation of large complex systems is not an easy task and requires both a good understanding of the simulated system and a good working knowledge of simulation techniques. This is illustrated by the following citation from *Newsweek*, March 18, 1991.

> In simulations before the attack, Air Force Computers projected that as many as 150 planes would be lost the first night. When the real thing came, all planes returned safely.

Fig. 1 gives a scheme for the simulation of a real life situation. Note that a simulation requires a mathematical simulation model. So simulation is not, as some might believe, a way to avoid using mathematics. Validation means to check if the simulation really is relevant for the practical situation which is to be simulated. This can be done, for example, by running the simulation program for a simple case of the practical situation for which an analytical solution is known.

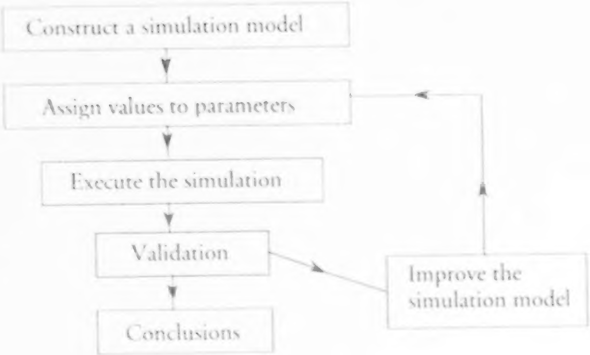

Construct a simulation model
Assign values to parameters
Execute the simulation
Validation
Conclusions
Improve the simulation model

Figure 1. *Flow chart for a simulation*

2. Teaching simulation

Elements of simulation can easily be covered in all basic stochastics courses. For example, when discussing various basic probability distributions, we should show not only their mathematical properties and areas of applications, but also how they can be simulated. Here the most effective simulation algorithms need not be used but methods which will give the students understanding and insight. Some examples are the following.

n = 60 p = 0.08 Number of observations = 1000
Mean = 4.877 Standard deviation = 2.16485

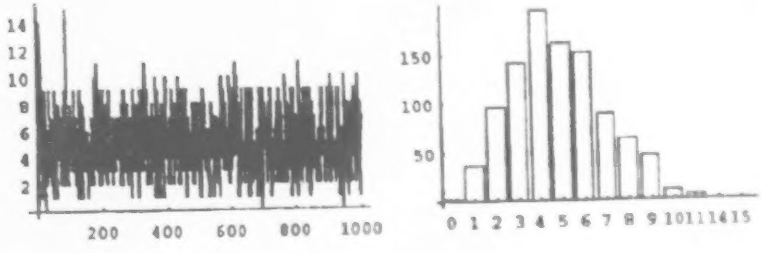


Figure 2. *Simulation of a binomial distribution*

1) The binomial distribution should be simulated by adding binary random digits. This will help the students see that a variable with a binomial distribution can be considered to be the sum of simple random variables, which can be used to calculate moments and to understand why a binomial distribution can be approximated by a normal distribution.

2) The Poisson distribution should be simulated through the simulation of the Poisson process by counting the number of events in an interval of fixed length. The Poisson process can be simulated by adding simulated observations of a random variable with an exponential distribution. This will help the students understand the important relationships between the exponential distribution, the Poisson process and the Poisson distribution.

3) The normal distribution should be simulated by adding a number of random numbers. This will give the students practical experience of the central limit theorem. A good exercise is to show why the addition of 12 random numbers is a good choice.

When using statistical computer software to do statistical data analysis, it is very appropriate to use simulated data produced by the computer.

Fig. 2 shows an example of results from such a simulation. In this case 1,000 observations on a binomially distributed random variable with n = 60 and p =0.08 have been simulated. The data can be interpreted as the observed numbers of defective units in samples of 60 from a production process with the failure probability p = 0.08.

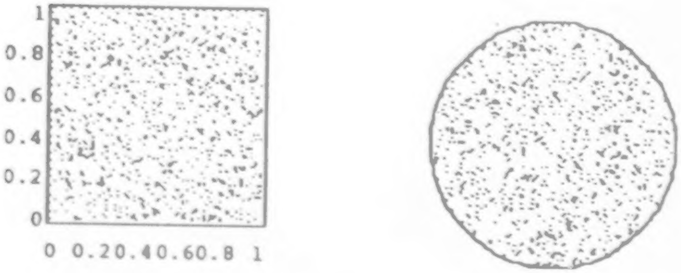


Figure 3. *Random points in square and circle*

It can also be instructive to have the students perform simulations of some simple random experiments. Examples of such experiments are:

1) Simulation of exponential or Weibull time to failure for some manufactured products.
2) Simulation of time to failure for series, parallel or other technical systems of components with random life times.
3) Generation of random points in a square and a circle.
4) Simulation study of the distance between two points chosen at random in a square.

For some engineering students simulation can be predicted to be such an important activity in their future profession that it is highly desirable to give them a specific course on simulation. Such a course should include at least one large practical simulation project. The following is a suggestion for the content of such a course.

1) Random numbers and random number generators.
2) Test of random number generators.
3) Simulation of simple random experiments.
4) Simulation of exponential, gamma, Weibull, normal, binomial, Poisson and hypergeometric distributions.
5) Simulation of the Poisson process.
6) General simulation by inversion, acceptance-rejection and the table methods.
7) Statistical analysis of simulated data.
8) How many observations?
9) Use of computer software.

It is not possible to discuss here details of how to treat these topics. It should, however, be pointed out that even if most statistics computer software has pre-programmed routines for the simulation of various standard distributions, it is important to discuss at least some algorithms for such simulations.

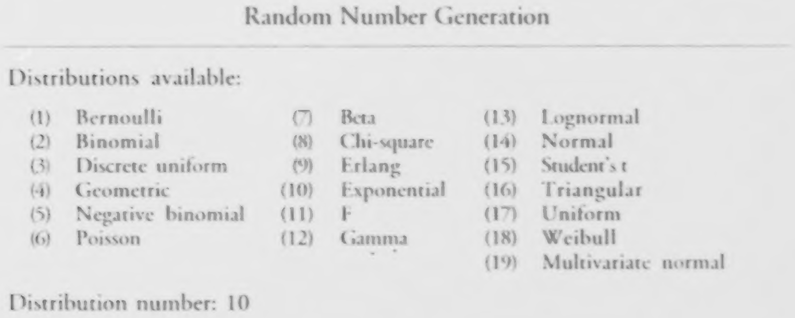
Textbooks on simulation often discusses at length several different methods for variance reduction. However, today the interest in such methods is not so great as it is usually very cheap to generate random numbers. It is probably enough that the students are informed, via illustrative examples, how to use control and antithetic variables for variance reduction.

3. Use of software

It is essential that the students when learning about simulation have access to computers and appropriate software. Most statistical computer program packages have commands for simulation. For example, Fig. 4 shows the menu for simulation in the Statgraphics system. Basic simulation exercises can also be made with general mathematical program packages like Mathematica. The illustrations in this paper have been made with this program. Simulation programs can also be written in general-purpose languages like FORTRAN, C, Pascal and BASIC.

There exist also special software for simulation and also for simulation of specific applications. This kind of simulation software can be used by the students for project work. In Law and Kelton (1991), where such simulation packages are discussed at length, the following programs are described:

GPSS (General-Purpose Simulation System)
SIMAN (SIMulation ANalysis)
SIMSCRIPT II.5
SLAM II (Simulation Language for Alternative Modeling)



```
                         Random Number Generation
-----------------------------------------------------------------------------
Distributions available:

  (1)  Bernoulli            (7)  Beta           (13)  Lognormal
  (2)  Binomial             (8)  Chi-square     (14)  Normal
  (3)  Discrete uniform     (9)  Erlang         (15)  Student's t
  (4)  Geometric           (10)  Exponential    (16)  Triangular
  (5)  Negative binomial   (11)  F              (17)  Uniform
  (6)  Poisson             (12)  Gamma          (18)  Weibull
                                                (19)  Multivariate normal

Distribution number: 10
```

Figure 4. *Menu for simulation in the Statgraphics system*

Bibliography

Hamming R.W. (1993), A Systems Engineering Approach to Education - Especially Mathematics, in D. Lum (ed.) *Proceedings of the 5th International Conference on Technology in Collegiate Mathematics*, Addison Wesley (to appear).

Law A.M. and Kelton W.D. (1991), *Simulation Modeling and Analysis*, McGraw-Hill, New York.

Ralstone A. and Reily E.D. (1983), *Encyclopedia of Computer Science and Engineering*, van Nostrand, New York.

Råde L. (1993), Probability and Mathematica, in D. Lum (ed.), *Proceedings of the 4th International Conference on Technology in Collegiate Mathematics*, Addison Wesley.

Råde L. (1993), Stochastic Geometry and Mathematica, in D. Lum (ed.), *Proceedings of the 5th International Conference on Technology in Collegiate Mathematics*, Addison Wesley (to appear).

Råde L. (1993), The art of Simulation, *European Journal of Engineering Education* (to appear).

OFF-CAMPUS STATISTICAL EXPERIENCES FOR UNDERGRADUATES

Carol Joyce Blumberg
Department of Mathematics and Statistics, Winona State University
Winona MN 55987-5838, USA

1. Introduction

Winona State University is a public liberal arts institution with approximately 7000 undergraduate students. Its roots are as a teacher training college and its strong ties to education continue to be important. The Statistics program is in the Department of Mathematics and Statistics. Officially, it is an option within the Mathematics major, but for ease of readability, for the remainder of this paper, the program will be referred to as the Statistics major. There are approximately 110 students with majors in the department, of which approximately 20 have chosen to pursue the Statistics major. About 2/3 of those choosing the Statistics major couple it with another major and/or a minor, although completion of a minor is not a requirement for the Statistics degree.

Students who decide to pursue a Statistics major must complete four quarters of calculus (5 credits each), a quarter of discrete mathematics, a quarter of introductory linear algebra, a course in foundations of mathematics, an introductory calculus based statistics course, a quarter of probability, two quarters of mathematical statistics, two quarters of applied statistics (analysis of variance and regression), a course in a computer programming language (usually Fortran) and a statistics project (for at least two hours of credit). All courses are four credit hours unless noted otherwise. The above listed courses account for 62 credit hours. In addition, the student must complete 20 hours of electives in Mathematics and/or Statistics. Most students choose a Practicum (Internship) experience as part of these 20 hours of electives. The Practicum experience used to be a required part of the major, but was dropped as a requirement three years ago to allow more flexibility in the major. The Statistics Projects and Practicum in Statistics experiences constitute the off-campus experiences that are the focus of this paper.

2. Overview of the Statistics Projects and Practicum experiences

The Statistics Projects experiences are offered on an individual or small group basis. This allows the faculty to tailor the projects better to the individual students' needs and to the opportunities available at the time. These experiences are usually ones where the students work on a project under the close supervision of a faculty member at the same time they are taking other courses.

The Practicum experience, on the other hand, is usually an off-campus experience where the students' daily supervision is done by an appropriate person at the Practicum site. A faculty member, however, checks with the supervisor and student at regular intervals. It is often completed as a temporary full-time job during a quarter when the student is taking no other courses or over a summer. Whenever possible, the faculty try to get the students paid for the work that they do when completing the Statistics Projects and Practicum in Statistics credits. This is not, however, always possible.

Both experiences require that the student work at least 25 hours for each credit hour elected. The students may elect between two and eight hours of credits for the Statistics Projects and between four and eight credit hours for the Practicum in Statistics. It is expected that the student write (or help to write) a report detailing a study that the student worked on for either the Statistics Projects or Practicum credits. This report usually includes an introduction, a brief literature review (where relevant), a methods of data collection section, a description of the analyses used, some important results, and a discussion of suggestions for the future. The requirements of the report are tailored somewhat to fit the individual study. In addition, they write a very short paper, of one to three pages in length, briefly describing the other activities they were involved in while completing their Statistics Projects and Practicum credits.

3. Short descriptions of some statistics projects

Survey for the friends of the Winona Public Schools

In the late summer and fall of 1988, two students and one of the faculty members became involved in what turned out to be one of the best Statistics Projects from the viewpoint of community relations and local press recognition. The students and faculty member, together with the district Superintendent, Directory of Community Education, and several parents, developed a form for use in a telephone survey. This survey had several diverse purposes that made the task of designing an unbiased survey very difficult. The students also helped to administer the survey, along with a group of parent and teacher volunteers. The students then analyzed the data. Finally, a report was written and disseminated to the media, school district, and other interested parties.

Water conditions in a backwater region of the Mississippi River

One of the students' mothers is very active in a group that helps the U. S. Fish and Wildlife Service collect data on water conditions at 18 sites in the backwaters of a certain area of the Mississippi River located near Winona. It turned out that a lot of data had been collected, but that no one in the Fish and Wildlife Service had the time to analyze the data properly. This student asked if she could analyze part of the data. The project appeared to be too big for one person, and so she worked together on it with another student. They prepared a report for the Fish and Wildlife Service detailing the results of their analyses.

Uses of the local lake

The local lake is a major attraction in town. It has slowly been transforming itself into a swamp. The local daily newspaper wanted to study the ways in which people use the lake and its surroundings and whether the public would pay the money needed to save the lake. They wanted to include the results of the study in a series of articles about the future of the lake. One of the Statistics majors helped the editors design a telephone survey instrument, determined how to get as unbiased a sample as possible of the users of the lake, and assisted in the data collection. At that point, the news editor left the newspaper. The study was dropped but the student was allowed to do some simple data analyses.

Use of "bike" paths around the local lake

At the beginning of the Spring of 1990, the chairperson of the Lake Winona Committee telephoned one of the faculty members. The Committee wanted to do an observational study of the ways in which the bike paths around the lake were being used. One of the first things that was done was to consult with the student who did the general telephone survey of the uses of the lake. He provided the two students who were to conduct this study with information about the use of the bike paths from his study. Based on his information and the expertise of a different faculty member, the students designed, carried out, and analyzed the data from this study. They wrote a report that was given to the Lake Winona Committee and then disseminated by that group to the various branches of the city government and to the local press.

Assisting quality control personnel

Four students worked on projects with quality control personnel from different companies. One student designed a series of interactive Minitab macros that built upon and adapted the existing Minitab commands (including control charts and other graphics) to Lake Center Industries' (a supplier for the automotive industry) needs for a particular set-up. Another student worked on a designed experiment involving some composite materials for ICI Fiberite (a composite materials manufacturer). A third student worked as a member of a quality control team at Watlow Industries (an electronic heating control supplier) that needed help with statistical analyses connected with a shelf-life problem. The experiences that these three students gained by having the opportunity to work closely with a quality control professional inside a plant were extremely worthwhile. The fourth student was already working as a full-time production worker at emd Associates (an electronic parts supplier). That company has been one of the leaders in the area of quality control in the Winona area. The student was allowed to design and run a designed experiment on one of the company's processes.

Customer Survey

One student, who was a double major in Statistics and Computer Science, helped design and pilot a mailed questionnaire dealing with customer satisfaction with the service they had received from IBM Rochester for a certain system and its associated products. She then developed a data entry program, entered the data, analyzed the data, and prepared a report. This student was already working part-time at IBM Rochester.

4. Short descriptions of some practicum experiences

Mayo Clinic

Over the past seven years, 12 students (at least one each summer) have completed their Practicum experiences by working as Interns in the Section of Biostatistics at the Mayo Clinic. Ten of these were summer interns, while two worked for one or more quarters during the academic year. As interns at the Mayo Clinic, the students receive intensive training sessions on the use of SAS as adapted by the Mayo Clinic. They are then either assigned to a single study/project or as a general data analyst. The general data analysts work on a variety of projects that are given to them by various full-time employees. One of the section of Biostatistics' master's degree statisticians supervises their day to day work and a Winona State faculty member only needs to check on things a few times a quarter.

The studies and projects that Winona State students have been involved in over the years are varied in terms of their fields of medicine. The students have been involved in all phases of studies from design through final analyses and report writing. The statistical techniques, both descriptive and inferential, that the students have used are as varied as the studies themselves.

Besides getting experience in real-world data analysis, the students are invited to attend the seminars and social events held by the Section of Biostatistics. The students also get the opportunity to work closely with students (both undergraduate and graduate) from several different universities. They also get the opportunity to talk to Master's degree and Ph.D. level statisticians about the graduate programs they attended.

Employee satisfaction surveys

Recently, a student, along with a faculty member, helped a group of employees from the Winona site of ICI Fiberite (a composite materials manufacturer) design an extensive survey of employee satisfaction. She then aided in the data collection, analyzed the data and together with this faculty member wrote a massive final report. The report they completed for the Winona ICI Fiberite operation was shown to several people at the Tempe, Arizona, site of ICI Fiberite. This faculty member along with this student and another student then completed the data analysis and final report for a very slightly modified version of the survey form from the Winona site that was used at the Tempe site. It also should be noted that two students, in two consecutive years, worked on similar surveys of employee satisfaction for another local corporation (emd Associates - an electronic parts manufacturer), but not as part of their Statistics Projects or Practicum credits.

Quality control practica

Three students chose to do their Practicum experiences as Quality Control interns at Watlow, Inc. (an electronic controls manufacturer). They assisted the Quality Control manager on a variety of projects. One student spent a major portion of her Practicum experience trying to figure out for which types of packages each of the various commercial delivery systems (e.g., Federal Express, U. S. Postal Service) was the most efficient in terms of delivery time and cost. The other two students were involved with various projects using a variety of Statistical Process Control (SPC) procedures, design of experiments, gauge reproducibility and repeatability studies, and capability analyses.

Another student worked for Lake Center Industries (an automotive industry supplier) a few hours a week over the course of an academic year helping to design and then analyze the data from a questionnaire concerning people's opinions of electronic heating control systems in automobiles. Another student began work several months ago as a full-time employee in quality control at Ashley Furniture. When hired, it was the company's intention to continue to employ her after graduation, not just for the Practicum experience. Recently, two students have completed full-time internships in quality control at emd Associates. These last three students have applied the concepts learned in the statistical process control and industrial design of experiments courses to a variety of processes.

Attitudes toward homework survey

One non-traditional student, who felt that she needed a Practicum experience that did not have set hours, decided to help the Curriculum Committee of the Winona Public Schools carry out a survey of the attitudes of parents, students, and teachers toward homework. The student, with only minimal supervision from the Curriculum Director and a faculty member, designed the three surveys (one for parents, one for students, and one for teachers), designed the methods of data collection for each group, analyzed the data, and prepared a written report for the school district.

5. Lessons learned

1. Care must be exercised continually that the students are doing meaningful activities. For example, it is important that students get some experience on data coding and data entry, but having them do only data entry and data coding is not a meaningful experience.

2. One must be very careful that the students do NOT spend too much time on their projects and Practica. The students get so involved in their projects that they sometimes do not put the effort needed into their other courses and their work quality in other courses is adversely affected.

3. Although the students are not happy about it at the time, the requirement of having to complete a scholarly written report is important. It gives the students practice in technical writing and gives the employer/recipient and Winona State University an excellent record of what the student did. Several students have commented one or more years after graduation that having to do these written reports was very helpful to them in graduate school or employment.

4. Finding good projects for weak students is hard. Also, the weak

students take more faculty time in terms of supervision than the stronger students. Finding off-campus Practica for these students has also been very hard. Allowing the students to find their own projects and Practicum sites helps. But, it is important that a faculty member still approve the projects and Practicum site.

5. It is important to have some form of written contract between the student and faculty member before the Statistics Project or Practicum is begun. The form used at Winona State is very simple. It has a projected date of completion, the number of credit hours being elected, and space for the student and faculty member to fill in responses to the following:

THE STUDENT AGREES TO:
THE FACULTY MEMBER AGREES TO:
METHOD OF EVALUATION:

Filling in such a form protects both the student and faculty member in case of a disagreement over the completion of the Project or Practicum.

6. Having two experiences, both a closely supervised one such as Statistics Projects, and a loosely supervised one, such as the Practicum, is the ideal. Every attempt should be made to have these experiences either build on one another or cover very different areas.

7. Although having two experiences is the ideal, having the Practicum in Statistics as optional rather than required has worked well. It allows the faculty the option of refusing to place weaker students in Practicum experiences and allows students pursuing a variety of goals to meet their goals in the way that is most meaningful to them.

8. Off-campus Projects and Practica are far superior to on-campus Projects and Practica for several reasons. First, they expose the students to the real-world work environment. The students also often form close working relationships with adults other than the Statistics faculty. It is our impression that the letters written by the off-campus contacts for these Projects and Practica are just as important, if not more important, to future employers and even to many graduate programs than those written by faculty members.

9. Supervising these Projects and Practica takes up a great deal of faculty time. It is time well spent. The faculty must be careful NOT to allow the Projects and Practica to take too much of their time and detract from their other activities.

10. The students should be free to choose which faculty member they want to supervise their Project or Practicum. Faculty members also should have the freedom, when necessary, to refuse to work with a student. There should, however, be a mechanism in place, either formal or informal, to make sure that every student can complete any Projects or Practicum credits required for the degree.

6. Acknowledgements

This paper is a shortened version, with some minor additions, of a paper of the same title presented at the 1992 Joint Statistical Meetings in Boston, MA and appearing in the *American Statistical Association 1992 Proceedings of the Statistical Education Section.* The author wishes to thank the students who have done Statistics Projects and Practica in Statistics over the last five years. Their wonderfully written reports were invaluable in the writing of this paper. All the projects and practica done by these students are not discussed in this paper. Apologies are given to the students whose projects and/or practica are not included. The choice of what to include was a difficult one. There were also some extremely valuable experiences done on-campus. Some of these experiences are described in the paper appearing in the *1992 Proceedings.*

THE EFFECT OF CONTEXT ON THE TEACHING OF STATISTICS AT FIRST YEAR UNIVERSITY LEVEL

Megan Clark
Institute of Statistics and Operations Research, Victoria University
PO Box 600 Wellington, New Zealand

1. Introduction

Clark (1993b) has described how at Victoria University in New Zealand there are two different but essentially equivalent (with respect to level of difficulty and statistical content) first year statistics courses. These courses are: STAT131 Data and Probability – the course recommended for students majoring in mathematics, physics, chemistry, computer science and engineering; and STAT193 Statistics for the Natural and Social Sciences – suggested for those majoring in biological sciences, social sciences and medicine.

The proportion of females in STAT193 is always higher than the proportion of females in STAT131. This is unsurprising, as approximately one and a half times as many females as males study biology in their final year of high school (Ministry of Education, 1992). Less predictable is that it is easier for female students to succeed in the STAT193 course than in the STAT131 course. As shown in Table 1, they consistently take a larger proportion of the A grades than their proportion of the class while the converse is sometimes true for STAT131.

Table 1. *Participation and success for female students in first year statistics courses*

	% female		female % of A grades	
year	131	193	131	193
1987	20	44	26	62
1988	30	36	27	48
1989	21	41	25	59
1990	37	52	37	57
1991	28	54	38	60
1992	29	64	20	73

The total numbers of students in these classes (220 and 480 respectively in 1993) together with the amount of available resources dictates the fairly traditional lecture format of the courses with a final timed examination.

2. The final examinations

Over a three year period some patterns have emerged. In the STAT131 final examination, male students often perform better overall. The examination as a whole is usually somewhat abstract. Any non-abstract questions often (> 50%) involve machinery, manufacturing or money and it is these latter questions in which males, in general, perform better (and are more likely to choose). The questions that female students tend to prefer in this paper are questions involving standard theory or algorithms. No questions or part questions over this three year period came from what one might term the "world of women" although in 1992 'she' was used as the pronoun in two general questions.

In the STAT193 examination male and female overall scores were either the same or significantly in favour of females. There are always two or three very abstract questions, usually about finite populations, the central limit theorem, power etc., but these questions are shunned by male and female students alike, and performance is poor in them. The nonabstract questions were predominantly 'people' or 'animal' questions and sometimes (11%) related to manufacturing. In this paper males and females attempt the same questions. The most frequently attempted questions were those involving people and the environment, for example: analyses of variance on scores in psychology tests, and abilities of mentally retarded people to learn basic life skills; contingency table analyses on spouse abuse versus family type, and on vaccination levels and ethnicity; a confidence interval on foetal heart rates, a t-test on chemical residues in birds. These were the questions in which the female students did particularly well.

The majority of male students in STAT193 are not studying in technical or mathematical areas. The problems that appeal to the females in the class are, by and large, the problems that appeal to these males who have mostly rejected or not been successful in mainstream mathematics. Work with secondary school pupils in New Zealand has confirmed a liking for problems in which students can see the connection between the statistics and their lives (Purser and Wily, 1992).

The numbers of Maori (indigenous New Zealanders) in these classes are too small to draw any statistically significant conclusions but some indications exist. A regression question relating pre-historic cave art in France to incidence of mammoth bones was chosen by a greater proportion of Maori students than European students and some Maori students did very well in this question. It seems reasonable that there are questions that are significant and meaningful to various ethnic groups and which enable them to demonstrate their knowledge. Furthermore, of five undergraduate first year mathematics or statistics courses offered, STAT193 was the only one to have more Maori female students than Maori males.

3. Course work

Success in the final examination, however, is not entirely due to the content of the examination. The courses themselves have differences. The nature of the weekly assignment problems in these courses differs. All assignment problems for three years were categorised into three groups: firstly, abstract and traditional (urns, dice, $P(A \cup B)$, coins, ...); secondly, problems about people and animals (blood pressure, bacteria, ...); and a category of non-traditional and non-abstract problems whose content could be seen as belonging to that sphere of activity that is traditionally thought to be appropriate to males (Morse code, football matches, the stock market). Over the three years STAT131 has had up to a quarter of its assigned problems from the 'boys world' and close to a quarter framed in terms of people or animals. STAT193, however, has had less than one eighth of its problems from the 'boys world' and between fifty and eighty percent have a people or animal context; i.e. there is far more emphasis on people problems in STAT193 throughout the course.

This suggests some reasons why female students might be preferring STAT193 as there is much research to show that "the most influential factor in girls attitudes to study is whether the subject is perceived as a 'male' or 'female' subject" (The Royal Society, 1986).

Another difference is in the staffing of these courses. STAT131 has three male lecturers while STAT193 has two male lecturers and one female lecturer. An exploratory study of 102 STAT193 students in 1993 strongly suggests (89%) that the students themselves do not believe this to be a factor.

4. The exploratory study

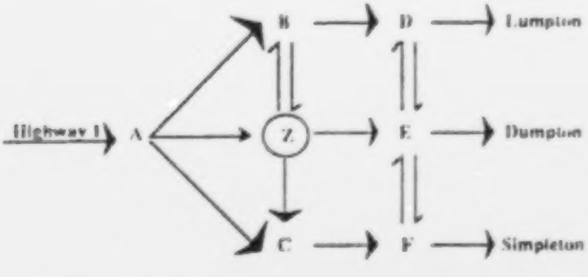
In 1993 a sample of 102 class members participated in an exploratory study which consisted of ten pairs of statistical questions. For each pair students were asked to specify which of the two options they felt that they would prefer to answer if they were required to or if they had no preference. They were also asked if they could say why they chose particular options.

This study clearly indicated that in some areas the students had very definite preferences (70% or more of the group) for particular contexts over others. In particular, problems involving money, the stock market, investments and any allied topic along with engine parts and other machinery oriented problems were soundly rejected by this group of students. This in part reflects the bias of these particular students who are not taking commerce courses or engineering but the strength of the reaction was a little surprising. There were further indications that no matter how relevant a question seemed on animal or botanical issues, if the question was in competition with a question involving people, then the people question would be preferred. This was highlighted when a choice between a question on soil types and one on blood glucose levels in diabetics was offered. Of those students that had a clear preference, twice as many opted for the question involving the diabetics. This was further illustrated by a question in which they had to choose between the contexts of a political opinion poll and the incidence of sheep intestinal cancer in New Zealand. These are both current and relevant issues. Forty three of the respondents would rather attempt the opinion poll question, forty three were neutral and only sixteen opted for the intestinal cancer. These results cannot be entirely explained by the relative proportions of social science to biology majors in the group (2:1). For those students able to explain their preference the criteria were clear and very similar. Those preferring the sheep cancer stated: "it's more interesting", "I'm interested in biology not politics", and "I'm not going to vote Alliance". Reasons given for preferring the opinion poll option were: "I can relate to it" (2), "I can vote so it affects me", "it's a current event" and "option B is gross". This also serves as a useful reminder that 'relevance' must not be the only consideration. Topics such as the destruction of the ozone layer may add relevance to our classes but we should be sparing in the use of such material so that statistics doesn't become irretrievably associated with the unpleasant (Clark, 1990).

This study confirmed results found by Purser and Wily (1992) with high school students. In particular that no matter how interesting an example is, if it is in competition with a standard, recognisable problem then the standard problem is preferred. The example used in the study is shown below. Students had to say whether they would prefer to do problem A or problem B or whether they had no particular preference.

A. The town of Muddleton is well known for its complicated oneway streets and complete lack of signposting. A diagram of the main routes through Muddleton is shown below.

Mr V approaches Muddleton along Highway 1, and is aiming to leave it along the road to Simpleton. Each time he approaches a cross road his wife randomly selects (with equal probability) any one of the choices open to him (U-turns are not permitted). Thus, approaching S, with probability 1/3 he takes the roads to B,Z,C. Approaching B from A, he goes to D or Z with probability 1/2 each, etc.



(⟶ means 1-way and ⇄ means a 2-way street)

What is the sample space associated with the "experiment" of driving through Muddleton? List out all possible points (i.e. routes) and find the probability associated with each.

B. An urn contains 4 balls coloured black (B), white (W), red (R) and green (G) respectively. Three balls are drawn, one at a time, without replacement, and their colours noted in order of selection.
Twelve of the simple events for this experiment are:
(B,W,R), (B,W,G), (B,R,W), (B,R,G), (B,G,R), (B,G,W), (W,B,R), (W,B,G), (W,R,B), (W,R,G), (W,G,B), (W,G,R).

By using trees or otherwise, list the remaining simple events and their associated probabilities.

Seventy-one of the 101 respondents to this question were clear that they preferred option B, sixteen had no preference and only fourteen expressed a wish to do option A. Many of the comments were quite revealing: "I like the urn problem, it reminds me of junior school", "seems simpler", "more like what's been taught", "more concise", "clearer", "easier", "less problem solving", "looks familiar", "tree diagram process is much more familiar than following the highway and road diagram", "cliched".

However one of the few Maori students in the class when opting for A stated: "related to everyday life" and another student stated that: "B is pointless".

This is not a simple situation however. It might be argued that the male driver and topic of vehicle routes would deter many female students from option A but the proportions of males and females choosing option B (and the other options) were not significantly different. Furthermore, this problem is rather typical of many spurious problems that ostensibly provide a relevant context. Option A is presented with a diagram, which to a few (4) seemed helpful but to some other students (4) in this class it was clearly perceived as a hindrance and reasons given for not preferring this option were:
"diagram"; "less words and I wouldn't need to take time in figuring out a diagram"; "look at the diagrams!".

A number of students (8) indicated that the comparative brevity of option B was the deciding factor.

However, the problems that students prefer as test items are not necessarily the same as the illustrative examples that they would prefer in class or the ones that they actually opt to attempt as was seen in the discussion of student choice in the final examination.

The indications at secondary school are that female students are doing better at problems that have "meaning and significance" (Howson and Mellin-Olsen, 1986) for them (Clark, 1993a). Purser and Wily (1992) in their work with high school students comment on "how much the students, particularly girls, were influenced by the gender orientation of the context of the question". We should probably proceed on the assumption that the situation is no different for university students. The exploratory study suggests that for straight from school students, the young women are less likely to have no preference among the options than young males tending to confirm the suggestion that context matters more to them ($t_{67} = 1.988$). This implies that we should present problems that make all groups feel at home in their statistics classes.

When the students were asked which they would prefer of an analysis of variance involving psychological test scores or one involving strengths of concrete mixes, 41% chose the test scores, 47% were neutral and only 12% opted for the concrete mixes. This was consistent with results from 1991 when students were required to choose one of these two as an assignment problem. Then 64% chose the psychological test scores and 36% the concrete mixes. In 1991, women were more likely to take the second option if the first was about concrete and overall 73% females took the psychological test score option and 53% of the males. In 1993, of the non-neutral responses, 74% of females and 84% of the males chose the test score option while males were more likely to be neutral (51% versus 44%) than females.

5. Assessment types

Not a great deal of work has been done so far with student preference and performance in different assessment modes. Until this year the STAT131 course has had no internally assessed component. However, the STAT193 course has a large internally assessed component: the Data Assignments. These consist of five linked assignments throughout the year all performed on a single data set. These data sets are specific to an individual student and consist of data on a set of hypothetical mothers and newborns from a region in a year and includes information on gender, weights of mother and baby, mothers' smoking habits and socio-economic group. Throughout the year the students investigate this data set. While some specific questions have to be answered these assignments are much more open ended than a traditional assignment, the student has much more control over what they decide to do, the material of the assignments is all ostensibly within a relevant social context and collaboration on the work is possible and appropriate up to a certain point. Furthermore, the results of these investigations have to be written up clearly. Thus these assignments can be seen as not as competitive as the final examination questions and much less threatening. If a student gets a very low score in one of these, there are four more in which to recover a respectable grade. These qualities are all qualities that the literature (Ministry of Education, 1992; Everybody Counts, 1989) describes as "girl friendly" and this proves to be the case. In STAT193 females do significantly better than males in both the traditional timed examination questions and in these Data Assignments, but they do 'more better' in this investigational, open ended setting as shown in Fig. 1.

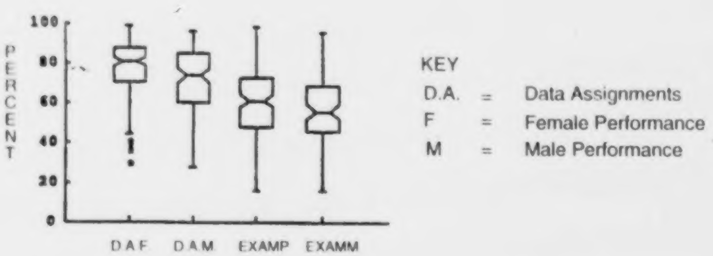


Figure 1. *Performance in examination and data assignments by gender*

The language of examples and problems makes problems more or less accessible to the student. Given the problem in terms of people, animals and everyday reality, female students and non-technical males appear to achieve better than when attempting the same material couched in the abstract or in terms traditionally associated with industry and manufacturing.

Underachieving in mathematics has been associated with female students for a long time and they may now perceive being good at mathematics as being unfeminine (Sherman, 1987). Now statistics is not mathematics, and there does seem to be a difference in its perception by young women, at least in New Zealand (Forbes *et al.*, 1990). We must be careful not to lose this advantage by mathematicising statistics unnecessarily and too soon. In particular we should be holding on to the data for as long as we can. Some of the ways we currently ask questions and assess work may be implicitly asking young women to opt out of belonging to the group of feminine women - an untenable demand for any educator to make. It is important to set work that supports the notion that statistics is a discipline that is entirely appropriate for women.

A characteristic of female students is their avoidance of risk taking (Murphy, 1980). Lorcher (1989) describes how economically and ethnically disadvantaged girls in Germany were less inclined to attempt questions than boys and middle class white girls. Avoidance of risk taking appears to be taken to extreme lengths in the former group of students. It is imperative that we make the situation appear to be less risky, something not done conspicuously well at the moment. One way to do this is to make the questions set for examination and in class more comfortable and less threatening than at present. Negative attitudes to study are consistently related to poor academic performance (Watkins, 1982), thus university study in statistics needs to be a more comfortable and equitable experience for all students.

Since over 50% of New Zealanders in paid employment use statistics (Knight *et al.*, 1993) - it is vital that statistics is made as accessible as possible to all sectors of the population.

Bibliography

Clark M.J. (1990), Applications: Handle With Care, *Teaching Mathematics and its Applications*, Vol. 9, No. 1.

Clark M.J. (1993a), Friend or Foe: The Pressures for Change, *New Zealand Journal of Mathematics*, Vol. 22, pp. 29-42.

Clark M.J. (1993b), Context and Content of Statistical Problems, *I.A.S.E. Newsletter supplement to Teaching Statistics*, Vol. 15, No. 1 Spring (to appear).

Everybody Counts (1989), *A Report to the Nation on the Future of Mathematics Education*, National Research Council, National Academy Press, Washington, D.C., USA.

Forbes S., Blithe T., Clark M. and Robinson E. (1990), *Mathematics for All*, Report of a research project for the Ministry of Education, Wellington, New Zealand.

Howson A.G. and Mellin-Olsen S. (1986), Social norms and external evaluation, in B. Christiansen, A.G. Howson and M. Otte (Eds.), *Perspectives on mathematics education*, D. Reidel, Holland.

Knight G., Arnold G., Carter M., Kelly P. and Thornley G. (1993), *The Mathematical Needs of New Zealand School Leavers*, Research Report, Department of Mathematics, Massey University, Palmerston North, New Zealand.

Lorcher G.A. (1989), *Learning Mathematics in a Foreign Language in Mathematics*, Education and Society, Science and Technology Education Document Series No. 35, UNESCO, Paris.

Ministry of Education Research and Statistics Division (1992), *Provisional data for Education Statistics of New Zealand 1991*.

Murphy R.J.L. (1980), Sex Differences in G.C.E. examination entry, statistics and success rates, *Educational Studies*, Vol. 6, No. 2.

Purser P. and Wily H. (1992), *What Turns Them On: An Investigation of Motivation in the Responses of Students to Mathematics Questions*, Burnside High School, Christchurch, New Zealand.

Sherman B.F. (1987), A Cautionary Note, *Educational Studies in Mathematics*, Vol. 18, No. 3, pp. 325-326.

The Royal Society (1986), *Girls and Mathematics*, A Report by the Joint Mathematical Education Committee of the Royal Society and The Institute of Mathematics and its Applications, London.

Watkins D. (1982), Identifying the study process dimensions of Australian university students, *The Australian Journal of Education*, Vol. 26, No. 1, pp. 71-85.

THE EDUCATIONAL TRACKS OF "THE MAKING STATISTICS MORE EFFECTIVE IN SCHOOLS OF BUSINESS" CONFERENCES

John D. McKenzie, Jr.
Babson College, Babson Park, MA 02157-0310, USA

1. Introduction

Statistical instruction is, and has always been, an important component of both graduate and undergraduate business schools in the United States. Most of these schools require at least one course in applied statistics. In the middle of the 1980s a small group of faculty members felt that there was a need to examine the content and delivery of such courses (and the use of statistics) at their schools. These concerned faculty members organized a conference at the University of Chicago in 1986 to discuss these issues. Based upon the success of the first conference, there now have been seven additional annual "Making Statistics More Effective in Schools of Business" (MSMESB) conferences. The last conference was held at the University of Minnesota. From the first conference that included workshops on industry perspectives on the curriculum; the basic statistics course; quality control, quality and productivity improvement, and experimental design; statistical computing and graphics; forecasting and time-series analysis; and the interaction of statisticians with researchers in accounting, finance, marketing, production, and other areas in business schools, there has always been an educational track at these conferences. The 1993 teaching and curriculum track had workshops on learning styles, improving students' writing and oral presentation skills, cooperative learning, and field projects. In this report the author, who is a member of the conference's planning committee, describes and summarizes all of the educational tracks of these eight conferences. In addition, he discusses the impact of these conferences on statistical education at business schools. He mentions the numerous benefits from attending such conferences for any group of statisticians. Finally, he reports on the plans for the 1994 conference to be hosted by Rutgers University.

2. The first MSMESB conference

In the middle of the 1980s George Tiao, Harry Roberts and George

Easton of The University of Chicago recognized the need for a conference to discuss the role of statistics in schools of business. For, while such schools have almost always had one or more required applied statistics courses since they began in the late 1800s, Tiao, Roberts and Easton believed that these courses had become stagnant. The content of these courses had not really changed for close to 30 years; the standard delivery for a longer period of time. There were few textbooks that included real data or made serious use of statistical computing. In addition, many of the instructors of these courses had limited knowledge of how statistics was used in their school's curriculum. Nor did they know much about the use of statistics in the workplace.

Hence they organized the first "Making Statistics More Effective in Schools of Business" (MSMESB) Conference at The University of Chicago on 20-21 June 1986. About 130 people interested in business statistics participated in this conference. Approximately one quarter of these attendees came from industry. The conference had four components: three plenary sessions, six task force workshops, reports from the workshops and discussions of these reports and follow-up. In preparation for the conference information was gathered to assist some of the workshop participants. For example, there was a comprehensive survey of the teaching of statistics in business schools (see Rose *et al.*, 1988). Except for one traditional presentation, the conference was organized to allow for active participation by all of the attendees.

The three cochairs produced proceedings of this first conference (Easton *et al.*, 1987). It, and subsequent proceedings of the conference, are available from George Tiao at the University of Chicago.

It was also decided that based upon the enthusiastic response from the participants that similar conferences would be held the next three or four years. Yet, this was not to be, because MSMESB conferences are now an annual event.

3. The next seven conferences

The second conference was held at New York University in 1987 (see Chatterjee and Tenenbein, 1988). It was followed by two Midwest conferences held at The University of Wisconsin (Finster *et al.*, 1989) and The University of Michigan (see Spivey, 1990). The 1990 conference was held at the Wharton School at The University of Pennsylvania (see Hildebrand, 1992). The University of Kansas was the site of the next conference (Hiller and Sherr, 1991). Approximately 120 people attended each of these conferences.

In 1992 the seventh conference was held at The University of Tennessee (Lin and Sylwester, 1993). The University of Minnesota hosted the most recent MSMESB conference (Benson, 1993). Over 175 people attended each of these last two conferences. Participants at the most recent conferences can broken up into two groups: regular attendees and first-time attendees. Many of the first-time attendees reside near the conference site. Making the conferences available to business school statisticians in different regions of the country and at a reasonable cost are two of the objectives of the MSMESB Planning Committee. This is an informally organized group of individuals who plan the conferences. The Committee meets only twice a year, once at the conference. It selects a chair for an upcoming conference and suggests some topics for that chair's conference.

Based upon what has been successful in the past, there is now an almost standard format for these two-day conferences. There are usually three tracks: one dealing with education, one dealing with quality and productivity and one dealing with an application of statistics to a major field of management. For each of these tracks there is usually a plenary session and a series of workshops, where the attendees actively participate in the program. A lot of time is intentionally scheduled between sessions for informal discussion. Three other successful additions to recent conferences have been plant tours before the formal conference, roundtable discussions during one of the luncheons, and contributed paper sessions, that enable some statisticians to attend the conference. One recent complaint (?) has been that the conference contains too many interesting activities in its two-day format.

The planning committee has recently established a vision statement, a mission statement and goals for the conferences. It is recognized that these may be altered in the future based upon the need for continuous improvement.

There has only been limited financial support for the MSMESB conferences. For, while the conferences did obtain a major grant from the Allied Signal Foundation, most of its support has come from the host institutions, including the time of the Planning Committee members.

One of the goals of the conference is to produce proceedings. In addition, Planning Committee members have presented summary reports of the conferences at the Joint Statistical Meetings and at annual meetings of the Decision Science Institute.

4. The educational tracks of first eight conferences

For all of the conferences there has been an attempt to assist the participant in teaching statistics and statistical thinking, where such thinking is based upon David S. Moore's ICOTS II formulation:

> The idea of a *process*.
> The omnipresence of *variation* in processes.
> Variation can be *explained*:
>   Statistical *control*;
>   *Randomness* and *distributions*;
>   Systematic effects – *regression*, etc.
> The need for *data* about processes.

One of the ways this has been done is by the presence of an educational track at each conference.

Here is a listing of the sessions related to pedagogy from the first eight conferences:

1986 Conference at The University of Chicago:
- What Business School Graduates Need to Know about Statistics: the Management Perspective;
- Making Statistics More Effective in Schools of Business: the Dean's Perspective;
- Transformation of the American Style of Teaching Statistics;
- Curriculum: Industry Perspectives;
- Curriculum: Basic Statistics Courses.

1987 Conference at New York University:
- M.B.A. Panel: Importance of the Statistics Curriculum in Relation to Different Areas in the Business World;
- Essential Changes for the Teaching of Statistics;
- Basic Statistics Course Curriculum: Report of the Curriculum Committee;
- Statistical Methods for Quality and Productivity Improvement: Designing a Course;
- Statistical Computing in Business Teaching.

1988 Conference at The University of Wisconsin:
- Quality in Education;
- Teaching with Emphasis on the Improvement of Delivery: There is More to a Course than Content;
- Teaching with Emphasis on the Improvement of Delivery: Pareto Analysis of Delivery Problems;
- Teaching with Emphasis on the Improvement of Delivery: Short Presentations on Successes and Failures in Dealing with Delivery Problems and Showing of Videotapes.

1989 Conference at The University of Michigan:
- Discussion of the Impact of PCs on Delivery in the Introductory Statistics Course;
- A Demonstration of the Classroom Use of a Macintosh Computer;
- The Role of Group Projects in Teaching Statistics;
- The Use of In-Class Demonstrations in Teaching.

1990 Conference at The University of Kansas:
- Teaching Managers in a Short Time: A Panel Discussion;
- The Bead Game: A Demonstration;
- Teaching Statistical Thinking to Managers;
- Teaching with the Case Method: A Demonstration;
- Workshop on Ph.D. Programs.

1991 Conference at The University of Pennsylvania:
- A Primer on Statistics, Quality, and Past Conferences Statistics, Total Quality Management and Higher Education;
- Improving Statistics Textbooks and Materials: What is the Process?
- Improving Statistics Textbooks and Materials: The Voice of the Market;
- Improving Statistics Textbooks and Materials: Roundtable Discussion - How Can the Process and the Product Be Improved?

1992 Conference at The University of Tennessee:
- Can TQM Be Applied to Teaching Itself?
- What We Think We're Learning from the Teaching Lab;
- Some Basic Theory about Sampling;
- Analytical versus Enumerative Issues;
- Industry Experience: Enriching the Education of Statistics Teachers;
- Using Live Data in Business Statistics Courses;
- Removing a Barrier to Motivation: Grades;
- Management Education;
- Texts & Other Information Modalities;
- Teamwork: Among Students and Among Departments in the MBA Curriculum.

1993 Conference at The University of Minnesota:
- Unlocking the Genius in Your Classroom: Tapping the Power of Learning Styles;

Improving Students' Writing and Oral Presentation Skills;
Active Learning: Cooperation in the College Classroom;
Live from the Classroom: Cooperation Learning in Action;
Minnesota's MBA Field Project Course;
Grass Roots Total Quality Management for Education;
Benchmarking Management Education.

Further information about each of these sessions may be found in the proceedings from each conference.

Many of these early sessions dealt with the appropriate curriculum for today's business statistics. There was a sense from the participants that less time should be spent on mathematical probability and hypothesis testing, and more time on descriptive statistics, time series, quality and productivity topics, sampling and statistical communication. Hence it not surprising that quality topics have been discussed in approximately one-eighth of the sessions.

Topics stressed in other sessions include improving the textbooks and statistical software in our courses and our delivery of business statistics courses. One of the major themes of these sixteen delivery-oriented sessions has been the importance of using teams both in and outside of the classroom. Most of these sessions would appeal to any instructor of applied statistics, not just instructors of business statistics.

5. Impact on business statistics

Attendance at one or more of the MSMESB conferences has definitely had an impact on what is taught in business statistics courses and the way it is taught. It is my belief that anyone who has attended a conference has come away with numerous ideas, many of which have been implemented in the classroom. For example, changes from the traditional, but ineffective, lecture method. Participants have also discovered links between statistics and other functional areas in business and became aware of how statistics is used by our students in the workplace.

It has also had an impact on the business statistics textbooks. Some new texts have been developed based upon the recommendations of the conferences. For example, Cryer and Miller (1991). Many authors have incorporated ideas from the conferences in their texts. The same is true of some software developers. Most texts, and some software packages, now include files with realistic business data.

Still there has not been as much change as one might expect. Here are two reasons for this lack of change. First, attendance at one or more conferences encourages participants to implement change, but only a limited number of business school statisticians have attended the conferences. Second, and most important, is the strong resistance to change present among all instructors, including instructors from schools of business.

6. Conference benefits

As explained above, there are numerous pedagogical reasons for business school statisticians to attend a MSMESB conference. Among the most important are being exposed to possible improvements in the curricula of business school statistics courses and in the delivery of such courses.

Still the primary benefit from attending a MSMESB conference is the ability to learn from each other. For example, to meet peers from other institutions who share similar educational experiences with you. The networking found at these vital conferences is superb.

7. Future plans

The next MSMESB conference is scheduled to be held at Rutgers University, The State University of New Jersey, on 9-11 June 1994. The chair of the ninth conference is Douglas Jones.

Based upon the success of the previous conferences, there are now many schools interested in hosting future conferences. Among the states being considered for a future conference are California, Alaska, Texas and Iowa.

It is expected that the 1994 and subsequent conferences will include an educational track. Hence there should be a benefit to any statistical educator who attends one of these conferences because many of the topics covered at the MSMESB conferences are not necessarily just for business school statisticians. Finally, attendance at one of these conferences may encourage statisticians from other disciplines to begin similar gatherings.

Bibliography

Benson G. (1993), *Making Statistics More Effective in Schools of Business: Report of the 1993 Conference*, The University of Minnesota, Minneapolis (to appear).

Chatterjee S. and Tenenbein A (1988), *Making Statistics More Effective in Schools of Business: Report of the Second Annual Conference, June 19-20, 1987*, New York University, New York.

Cryer J.D. and Miller R.B. (1991), *Statistics for Business: Data Analysis and Modelling*, PWS-KENT Publishing Company, Boston.

Easton G., Roberts H.V. and Tiao G.C. (1987), *Making Statistics More Effective in Schools of Business: Report of June 1986 Conference*, The University of Chicago Statistics Research Center, Chicago.

Finster M.P., Frees E.W. and Miller R.B. (1989), *Making Statistics More Effective in Schools of Business: Report of the Third Annual Conference, June 24-25, 1988*, The University of Wisconsin, Madison.

Hildebrand D.K. (1992), *Making Statistics More Effective in Schools of Business: 6th Annual Conference Transcription of Sessions*, The University of Pennsylvania, Philadelphia.

Hiller S. and Sherr L.A. (1991), *Making Statistics More Effective in Schools of Business: Report of the Fifth Annual Conference, May 31-June 2, 1990*, The University of Kansas, Lawrence.

Lin D. and Sylwester D. (1993), *Making Statistics More Effective in Schools of Business: Report of the 1992 Conference*, The University of Tennessee, Knoxville (to appear).

Rose E.L., Machak J.A. and Spivey W.A. (1988), A Survey of the Teaching of Statistics in M.B.A. Programs, *Journal of Business and Economic Statistics*, American Statistical Association, Alexandria, Virginia, pp. 273-282.

Spivey A. (1990), *Proceedings of the Conference on Statistics Applied to Marketing Management*, The University of Michigan, Ann Arbor.

STATISTICAL EDUCATION IN DEVELOPING COUNTRIES: THE PROBLEM OF 'MARKETING' UNIVERSITY GRADUATES

Hristo Nikolov*, Erica Keogh* and Karl Stielau**
**Department of Statistics, University of Zimbabwe*
P.O. Box MP 167, Mount Pleasant, Harare, Zimbabwe
***Department of Statistics & Biometry, University of Natal*
P.O. Box 375, Pietermaritzburg 3200, South Africa

1. Introduction

The focus in this paper is on the attempts of the Statistics Department at the University of Zimbabwe (UZ) to adequately equip its graduates to become useful contributors to society. The needs of a developing country such as Zimbabwe are of primary importance; a consequence is the careful design of syllabi and course structures to be as relevant as possible. We distinguish between general trends in Africa and the particular issues faced by UZ.

It is encouraging to note that in the last decade more attention has been paid to researching aspects of statistical education in African countries. Two recent examples are Fresen and Fresen (1990) and Fresen (1992). It is imperative that developing Africa maintains itself statistically during the coming decade.

2. Historical background

The challenges and problems experienced by the Statistics Department at the University of Zimbabwe are in one sense unique and in another sense part of a pattern evident in several developing African countries. A clearer understanding of these challenges and problems may be achieved by a brief overview of general trends, similarities and differences between tertiary institutions offering statistics education in developing African countries. A comprehensive survey of all institutions has not, as yet, been conducted; the intention here is to provide a context for the subsequent discussion.

Statistics courses are offered by virtually all universities in developing Africa though in some instances it is impossible to major in the subject or to pursue postgraduate studies. Statistics at such institutions is often seen as an area with which students should be reasonably familiar; no specialised knowledge is required, merely an ability to summarise data and perform simple inference. It is difficult to criticise this attitude, as the host country usually has more pressing needs.

The following description of a "typical" African university statistics department is a generalisation based on:

- a random sample of South African universities;
- the University of Namibia (UN);
- the National University of Lesotho (NUL);
- the University of Zimbabwe (UZ);
- the National University of Science and Technology (NUST) in Zimbabwe, as part of the Department of Applied Mathematics.
- the Anambra State University of Technology (ASUTECH) in Nigeria.

A total of seven Universities in the region were thus considered. We briefly describe below a statistics department which is as representative as possible of the above institutions.

- The institution is typically fairly new. Of the above mentioned institutions four (Vista in South Africa, UN, UZ and ASUTECH) have statistics departments which are less than ten years old. A consequence of newness is a necessary concentration on service course teaching, coupled with a difficulty in attracting staff with research interests. Offering postgraduate courses is commonly not feasible.

- In most cases statistics has been taught prior to the formation of a department by:

a) pure or applied mathematicians with relatively limited backgrounds in statistical theory; or

b) social scientists with interests in fairly restricted areas of statistics, such as community surveys and nonparametric statistics.

This again impacts on the ability of staff to arrange for higher studies to be pursued by students.

- In a typical Southern African university a high proportion of expatriate staff is to be expected. While this is not in itself wrong, it is clearly desirable that in the long term a majority of academic staff at such an institution should be nationals of the country in which it lies.

- At two of the institutions arrangements were in place for the sharing of teaching duties: staff of the University of Cape Town have assisted in teaching at the University of Namibia and staff of the University of Zimbabwe have given courses at NUST. This is a desirable development, but the high cost of travel inevitably leads to courses being taught intensively rather than at a more leisurely pace.

- In all institutions an encouraging development is detected: changes to school syllabi have led to a greater understanding, on the part of university entrants, of the nature of statistics. It is not as difficult to attract able students to the discipline as it was in the past. However, increasing student numbers have led to:

- difficulties in providing adequate facilities for the incorporation of computer work into the courses. Here the primary difficulty is in the provision of terminals for student use;

- increasing entry qualifications to programmes. All institutions use a school-leaving points system for entry decisions. When inadequate schooling is available the question of reliability as indicators of these points must be addressed.

Such then are the general features of a tertiary institution offering statistics courses in developing Africa. We now address as a particular case the University of Zimbabwe.

This paper is concerned with the 'Marketability' of University Graduates in Statistics, in Zimbabwe today. In order to substantiate our claims to this we first need to describe the structure and content of the university statistics programmes.

3. Production of graduates in Zimbabwe

3.1 *Degree structure*

The Statistics Department, University of Zimbabwe, is located in the Faculty of Science and thus our regular students, working towards BSc degrees, study science subjects. The Department accepted its first students in March 1989 and has since then produced two sets of graduates, in 1991 and 1992. Students, aged between 18-20 years, are accepted into the Statistics Department on the basis of their A-level results, where the A-level examinations are those of the University of Cambridge Examining Board, UK (Zimbabwe has been in the process of localising its O-level syllabi since independence in 1980, which localisation is now virtually complete, and hopes to move towards localisation of A-level syllabi as from 1995). Any student entering the University of Zimbabwe is required to have a minimum of two subjects passed at A-level with additional requirements for particular faculties. The Statistics Department has, up to now, required that a student is eligible for entry only if he/she has achieved a grade C in mathematics in these A-level examinations. Exceptions are often made, since the department is well aware that the schooling system is far from perfect and a student with an E in Mathematics at A-level, from a school with no mathematics teacher, is likely to be a 'better' student than one with an A from an established, well-to-do school, with modern facilities.

A first year student takes three subjects; in the case of a student taking Statistics I he/she is also required to take Mathematics I and then takes one additional subject, viz one of physics, chemistry, biology, biochemistry, computer science, geology, or geography.

Having successfully completed first year each student, usually, continues studies in two of the three subjects, which will then form the degree majors. A course unit system has recently been introduced in the Faculty of Science, enabling students to 'pick-up' courses in other disciplines, but the overall structure remains that of two majors. A student will study these two chosen subjects through second and third year and will then have qualified for the award of a BSc degree, majoring in those two subjects.

There are various options open in each subject which will entitle the better students to pursue an honours option. For statistics the option is available to those who perform particularly well in the first year; such students are invited to join the honours programme, in which case he/she will study statistics only, as a double subject, through second and third year, and will emerge with a BSc Honours degree in statistics. This same qualification is available as a fourth year Honours degree after obtaining the BSC general degree. The department offered the Part III/IV Honours programme for the first time in 1992.

The Department will be offering a one-year Diploma in Statistics for the first time in 1994 and hopes to enrol students for the MSc programme by 1995. A post-graduate research degree, the MPhil, is currently available but is hampered by the lack of stable staff for supervisory services.

3.2 *Course content*

In the first year of study in statistics a student receives a basic grounding in theoretical and applied statistics; thus the first year courses consist of Probability I, Inference I and Applied Statistics I and II. The use of a computer package (MINITAB in our case) in practical sessions enables students to work with 'real life' data sets, using standard exploratory techniques, including non-parametric methods and introductory regression.

In the second year of study we aim to prepare a student to deal with statistics today, specifically statistics in Zimbabwe. Further theoretical work is pursued, viz Probability Theory II and Inference II. Further computing is introduced via the statistical computing course which focuses on SAS but includes introductory sessions on DOS, the use of word-processing packages, data-bases and spreadsheets. Further statistical techniques are introduced via the course on regression analysis, including residual analysis. The relevance of Statistics in Zimbabwe today is introduced via the two courses:

a) Social and Economic Statistics;
b) Survey Methods.

In the first, Social and Economic Statistics, we endeavour to present an overview of micro-economics in the Zimbabwean context and an introduction to demographic techniques. Students are made aware of the importance of a governmental statistical agency, of the value of socio-economic indicators, of the uses of demographic data, etc.

In the second, Survey Methods, we require the class to conduct a survey, from beginning to end, including questionnaire design, sampling procedures, interviewing, and ethics, analysis of data collected and report writing. Topics surveyed in the past include Student Health Services, Library Use and Mis-Use, and Sports Participation. Each of these has had the student population as the study population. This year, 1993, we aim to move off campus and to interview residents in a high density suburb of Harare on matters of readership, i.e. 'What are People Reading in Zimbabwe Today?'.

In the third year of study we aim to equip a student with the statistical techniques he/she is likely to need in the real world of Zimbabwe. We offer a final course in inference, focusing on distribution-free and asymptotic methods. Further statistical techniques are presented via courses in Experimental Design, Time Series Analysis and Multivariate Methods, all of which are dependent on computer usage in practical sessions and one of which requires project work to be undertaken. Students are, in addition, required to study two courses from the following four, viz Official Statistics, Operations Research, Demography and Econometrics; unfortunately not all options are available every year due to staff shortages.

At the Honours (fourth year) level, a student selects courses from a variety of topics including Stochastic Processes, Risk Theory, Biological Assay and Modelling, Statistical Ecology, further Official Statistics, further Econometrics, Industrial Quality Control and others. In addition, a project has to be undertaken and here we require that a student identifies for him/herself a topic for investigation, so seeking to promote the growth of an inquisitive, enquiring approach as a statistician. Some examples of project proposals for this year (1993) are:

a) Modelling the Consumption Function for Rural and Urban Zimbabwe;

b) A Study of Electrical Load Forecasting in Zimbabwe;

c) A Retrospective Study of Cardiovascular Disease Mortality in Zimbabwe;

d) The Impact of AIDS on the Economically Active Population of Zimbabwe.

Each of the above seeks to explore a topical subject and is thus relevant to the people of Zimbabwe today.

At the Diploma Level the Department seeks to upgrade a graduate with a degree in an applied subject, such as economics, psychology, agriculture, business, etc., to the Honours level so as to prepare him/her to enter the MSc programme in Statistics at a later date. Thus students here will study a variety of courses selected from the second and third years and honours programmes, each according to his/her needs. For many such students the need will be for additional theoretical courses as well as further training in techniques such as survey methods, experimental design, multivariate methods, etc.

At the Masters level we hope that a student will be able to select courses concentrated in one of two streams, viz the economic, industrial, governmental, insurance stream or the health, agricultural, environmental stream. We aim to rely on subject experts to provide relevant background material to the students while we statisticians focus on relevant techniques within the subject matter. Thus for example in a course on epidemiology we will expect medical staff to provide a subject matter basis from which we can develop relevant methods for data collection, analysis, interpretation, and dissemination. The programme will be part-time, emphasising our wish that statistics is to be used daily and thus encouraging students to work with what they learn. The second half of the programme will consist of individual research, focusing upon a relevant topic in the field of interest.

4. 'Marketability' of graduates in Zimbabwe

4.1 *Expectations from students*

Now we can move to consider the future of our students. It is not only those who graduate with a major in statistics that we consider, since the student with only first year studies in statistics is also liable to be called upon to utilise his/her knowledge. As stated previously, we aim, in the first year, to provide a solid foundation in statistical theory and techniques. Many of our students will become school teachers; our first year content well prepares them to teach statistics to A-level and, in addition, provides them with experience with real data so as to make the classroom situation relevant and exciting. Similarly, for those who will find employment in the public and private sectors, the familiarity gained with statistical techniques of data analysis and computing prepares them adequately for applications in their field of employment.

We can consider our studies to degree level from three aspects, viz theoretical, technical and applied. The theoretical knowledge provided through the three years of study exposes students to all the well known notions of inference including non-parametric methods, sequential analysis and asymptotic methods. The technical knowledge provided covers in depth the techniques of regression, design of experiments, time series, sampling and survey methods and multivariate methods. The techniques presented, with the all-important practical computing experience, combine with the theoretical justification to equip a student to face the challenges of statistics in practice in various employment sectors.

It is in the applied courses offered that we aim to really ensure that our graduates are valued by employers in all sectors of the community. The course in social and economic statistics followed by those in econometrics and demography are aimed at a student seeking employment in the economic sector, including banking and insurance. Replacing demography by operations research aims for the commercial and industrial sectors while replacing econometrics by official statistics looks to the governmental sector. We aim to give students the confidence to develop further their basic knowledge so as to expand those particular skills required in the job situation. Their computing experience serves them well since many employers today demand computer literacy coupled with particular expertise in one or more fields. We encourage students to select the second major subjects with care and encourage them to market themselves using the complementary pair of majors as their background. Research oriented jobs in laboratories will thus become available to students with biology, chemistry or biochemistry complementing statistics. Industry will find places for those with mathematics, physics or geology as the second subject. Geography will secure a place with the development sector while computer science will very well suit the computer servicing sector.

Postgraduate studies in the department aim to enable a student to identify and pursue an area of specialisation in applied statistics. The Honours programme seeks to further develop the basic courses along the two streams mentioned previously. These students usually have no second major subject and thus we feel it is imperative that they be given a grounding in the applied area so as to equip them to adequately interpret results. Economically based studies continue with courses in stochastic processes, risk theory for potential actuarial studies, and further econometrics. Governmental applications are continued with further official statistics while industrial uses are explored in the quality control course. Specific to agriculture, health and the environment are the bioassay and modelling and the statistical ecology courses.

The Diploma studies are aimed at the better students from other disciplines; students who have obtained a first degree in the applied area, with some statistical content, and now wish to pursue in depth the statistical aspects of their specialisation. Such students thus require theoretical and technical support and the second and third year courses adequately provide what is required. Each student is assessed on an individual basis and provided with a programme suited to his/her specific needs.

By the time a student reaches Master's level we hope that he/she has selected a stream in which to concentrate his/her studies. We hope to be able to cater for the two streams of students offering a variety of courses in the specialist areas. We hope that these students will find employment at the management level, drawing on their specialist knowledge to direct research, to restructure production procedures, etc.

4.2 *Staff participation*

Ambitious aims as we have in the Statistics Department, UZ, depend for their fulfilment on staff cooperation and staff activities. At present the Department has members of staff, on contract terms, from 6 different countries. While we have a programme of staff development which necessitates, at present, the sending of elegible Zimbabweans out of the country for further training, we will nevertheless be dependent, for some years, on employing expatriate staff. We feel students can benefit from such visiting staff in the matter of experience in, exposure to and variety of, the statistical activities of those staff from different areas of the world. We encourage staff to seek opportunities to conduct research with prospective employers in various sectors of the community. For example the past year has seen staff involvement with the Central Statistical Office in the conduct of the National Census (1992) and in the production of an Environmental Statistics Handbook for Zimbabwe, with the Ministry of the Environment in the Monitoring of the Environmental and Socio-Economic Aspects of Illegal Gold Panning in Zimbabwe, with the Ministry of Social Welfare in the Monitoring of the Social Dimensions of Adjustment within the ESAP programme, with leaders in industry in training managers in the use of the 7 tools of statistical quality control, with agriculturalists in experimental design and analysis, with health experts in the monitoring of AIDS and other epidemics and the training of health workers in this area, etc. Free advice is constantly given to members of the University who require statistical support, including computing, in their research activities. By these activities we hope to illustrate to students the vast potential for statistical applications that awaits them on graduation.

5. Conclusion

We in the Statistics Department, University of Zimbabwe, feel we are designing a unique and innovative programme of studies for our students. In offering courses such as Official Statistics and Statistical Ecology we aim to show that developing Africa has the potential and the ability to properly prepare students for prospective employment. We have produced some 60 graduates in our short life time and thus it is too early to demonstrate that our aims and aspirations are effective. The economic recession further confounds our outlook since few employers are looking to expansion at this point in time.

Bibliography

Fresen J. (1990), *Technical Report*, University of the Western Cape.
Fresen J. and Fresen J. (1992), *Unpublished conference paper*, South African Statistical Association, Port Elizabeth.

UNIVERSITY LEVEL STATISTICAL EDUCATION IN HUNGARY

Jenö Reiczigel
Department of Biomathematics and Informatics
University of Veterinary Science
H-1078 Budapest, István út 2, Hungary

1. Introdution

During the past ten years, I have taught probability theory and statistics at several universities in Budapest. All these courses were for students in applied fields; sociology, business and economics, agriculture, etc. In all the courses, basic level probability theory and statistics were taught, but from different points of view and with different emphases.

This paper summarises my experiences, presenting the courses in a common framework, so that some comparisons may be drawn concerning their aims, their content, the teaching methods employed and the main difficulties encountered. This short review of university level statistical education in Hungary may give an opportunity for comparisons with educational programmes in other countries. My experiences have shown that the main problems and the general solutions in other countries are often very similar to ours. This fact may strengthen the cooperation and exchange of ideas between people working in the field of statistical education.

I have not included the education of statisticians in this report. In Hungary, there are two different ways to learn to be a statistician. One is in the faculty of mathematics, and the other is in the faculty of economics. Of course, the main emphases are different: a theoretical (mathematical) approach in the first context, and an empirical (data oriented) approach in the other one. A third possibility is just being organised in the faculty of sociology, where the education of survey statisticians is going to start this year.

The system of higher level education in Hungary is much less flexible than is usually the case in Western countries. This means that, although the system has started to become more flexible than it was in the past, students generally have far fewer choices concerning the lectures and seminars that they will attend.

The following aspects of each course are considered: the length, a summary of the material covered and its particular emphases, the use made of computers, the preceding preparatory courses in general mathematics and the examinations requirements.

2. Some explanatory notes about the institutions included

Table 1 lists the institutions covered by this report, with the abbreviations which will be used throughout the paper. An asterisk by an entry indicates that the author has had personal teaching experience at that institution.

- At the Technical University, there are several faculties with their own mathematical departments and students with various specialisations. Thus several different teaching programs exist. It is typical, however, that only a minimum of statistics is taught.
- The teaching programmes at universities in the field of agriculture are quite similar to that of the University of Horticulture and Food Industry which is included in this report.
- At the medical universities, only students of pharmacy are taught by the mathematics departments. Students of medical faculties learn both mathematics and statistics within the context of a biophysics course. In spite of the lack of precise data, it is certain that both the time and also the material covered are less than where specific courses in statistics were given.
- The two faculties of pharmacy are not considered separately, although at Semmelweiss University somewhat less time is devoted to statistics, compared to courses at Szent-Györgyi, and no computer knowledge related to the statistics course is required.

Table 1. *The institutions covered by this report*

University of Veterinary Science	VETER	*
Eötvös University, Faculty of Sociology	SOCIOL	*
University of Horticulture and Food Industry		
Faculties of Horticulture	AGR	*
Faculties of Food Industry	AGR.ENG	*
International Business School, Budapest (Oxford Brookes University)	BUSINESS	*
Budapest University of Economics	ECON	*
Semmelweiss Medical University, Budapest and Szent-Györgyi A. Medical University, Szeged		
Faculties of Medicine	MED	
Faculty of Pharmacy	PHARM	
Budapest Technical University, Faculty of Electronics	TECH	

3. Some quantitative data about the courses

These courses are part of the students' basic education, so they take place in their first or second year programmes. Typically, students end their studies of statistics with these courses, although in case of certain specializations at faculties of economics and sociology, students may (or must) attend additional advanced courses in statistics (e.g. on multivariate methods), but these are not taken into account here.

Table 2 gives a summary of the total number of lectures and seminars devoted to probability theory and statistics at the various faculties. At seminars, students typically deal with numerical examples, doing exercises that may involve either manual computations or the use of the computer. The numbers shown in Table 2 represent hours, the length of a semester being taken as typically 15 weeks. Partly or wholly faculty-based lectures or seminars are noted by an (F) appearing after the number.

At Budapest University of Economics, the data in Table 2 concern the basic courses, that is those courses which are obligatory for all students. In fact, for some specialisations, much more statistics is taught, and it is here, in these specialist contexts, that most of the more advanced statistical methods which are taught would occur.

At technical faculties, where a lot of mathematics is taught, it is difficult to decide how much of it is directly relevant to statistical education, and therefore how much of it should be considered here. This is the reason for the NA (Not Applicable) entries in Table 2.

Table 2. *Total amount of lectures and seminars on probability theory and statistics, and preparatory general mathematics at the various faculties* (*in hours*)

faculty	lectures	seminars	lectures	seminars
	in probab./stat.		in general maths	
ECON	75	105(F)	60	60(F)
BUSINESS	30	90	0	0
SOCIOL	60	30(F)	30	30(F)
AGR	30	60	30	30
AGR.ENG	45	30	60	60
TECH	30	30	NA	NA
PHARM	15	30	30	30
VETER	24	16	21	14

The information in Table 2 may be illustrated in the form of a scattergram as shown in Fig. 1. Lectures and seminars have been equally weighted and the scales on the two axes - statistics versus general mathematics - are not intended to be equivalent. It can be seen that the students receiving the most statistical education are those of the business and economics faculties. It may be surprising, however, to see that medical students seem to be the least receiving statistics education.

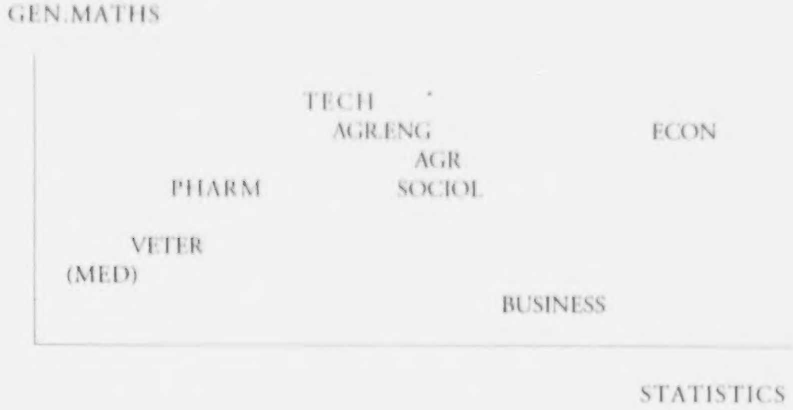


Figure 1. *Scattergram of the total amount of time devoted to statistics and general mathematics.*

4. The material covered

In listing the material of these courses, I will follow the usual grouping; probability theory, descriptive statistics and inferential statistics. In fact, except in the business and economics faculties, descripive statistics are of only marginal importance. In the business and economics faculties, however, descriptive methods and presentation techniques, graphics, etc., have particular importance, but strict probability-based theoretical foundations of inferential methods are often replaced by intuitive approaches. The teaching may be shared between the departments of statistics and mathematics, thereby providing both approaches.

In all faculties, statistics courses include some computer awareness, and the use of some statistical software as well. Different program packages are used, determined partly by traditional, and partly by financial factors. Besides computer usage, students are required to be able to carry out some basic statistical computations like a t-test or a chi-square test for independence, manually as well (although calculators can be used), and to be able to use probability tables.

From general mathematics, the following units can by regarded as preparatory with respect to statistics: set theory, matrix algebra, infinite series and sums, differentiation and integration of uni- and multivariate functions (of course, these are not explained in detail at all faculties).

Concerning probability distributions, it should be noted that those distributions used in statistics, like chi-square, Student-t, F, etc. are generally not discussed in the probability theory course, but rather they are left to the statistics course.

Table 3. *The material covered in the various faculties*
(L = learnt in detail; M = mentioned but not learnt in detail; - = not learnt)

	ECON	BUSINESS	SOCIOL	AGR	AGR ENG	TECH	PHARM	VETER
Probability theory								
axiomatic foundation of probability theory	M	-	M	-	M	M	-	-
random variables	L	L	L	L	L	-	M	L
univariate discrete probab. distributions								
uniform, binomial, Poisson	L	L	L	L	L	L	M	M
others	L	-	-	M	M	L	-	-
univariate continuous probab. distributions								
uniform, normal	L	L	L	L	L	L	L	L
exponential and others	L	L	-	-	-	L	-	-
sampling								
sampling with or without replacement	L	L	M	L	L	L	-	M
sampling distributions of random variables	L	L	-	-	-	M	M	-
limit theorems	M	-	-	-	M	M	-	-
relations between random variables								
joint distribution, condit. distribution	L	-	M	M	M	M	-	-
association	L	-	L	-	-	-	-	-
correlation	L	L	L	L	L	L	L	L
multivariate distributions								
(multinomial, normal, etc.)	L	-	-	-	M	L	-	-

Table 3. *The material covered in the various faculties* (continued)

	ECON	BUSINESS	SOCIOL	AGR	AGR ENG	TECH	PHARM	VETER
Descriptive statistics								
tabular and graphical presentations of data	L	L	L	M	M	M	M	M
quantitative measurements	L	L	M	M	M	M	M	M
sampling methods	L	M	M	-	-	-	-	-
index numbers	L	L	-	-	-	-	-	-
Inferential statistics								
theoretical considerations of								
estimation	L	M	-	-	M	M	-	-
hypothesis testing	L	-	M	-	M	M	-	-
interval estimation for the								
probability (proportion) or mean	L	L	M	L	L	L	L	L
variance	L	L	-	L	L	L	-	-
one- and two sample tests for the								
probability (proportion) or mean	L	L	L	L	L	L	L	L
variance	L	-	L	L	L	L	L	M
chi-square tests of goodness of fit and independence	L	L	L	L	L	L	L	L
other nonparametric tests	L	-	M	-	-	-	L	-
regression analysis								
linear with one independent var.	L	L	-	L	L	L	L	L
nonlinear with one independent var.	L	-	-	L	L	-	-	M
multiple linear regression	L	-	-	L	L	-	-	-
ANOVA								
one-way	L	-	-	L	L	-	-	L
multiway	L	-	-	L	M	-	-	-
experimental designs	L	-	-	L	-	-	-	-

In more advanced courses, in later studies at the faculties of economics and sociology, essentially all advanced statistical methods occur but mostly with little respect to their mathematical basis, especially if they are discussed in the so called empirical data analysis courses.

5. Examination requirements

The assessment requirements differ in the various faculties. A summary is given in Table 4. In the written/oral column, a W+O may have the meaning that the course is ending with an oral examination but students are required to work out exercises during the course. Where computer knowledge is not indicated, it may be included in another, not strictly statistical course (e.g. business information systems, empirical data analysis, etc.). Parentheses in Table 4 mean that those requirements are of secondary importance for some reason (e.g. an oral examination at the faculty of sociology is only for those who failed to pass the written examination). Written examinations take from 1.5 to 3 hours working time. Theoretical questions in written examinations may be of either multiple-choice or essay type.

Table 4. *Examination requirements*

faculty	written/oral	theor./exerc.	manual/computer
ECON	W+O	T+E	M
BUSINESS	W	(T+)E	M
SOCIOL	W(+O)	(T+)E	M
AGR	W+O	T+E	M+C
AGR.ENG	W+O	T+E	M+C
TECH	W	E	M
PHARM	W+O	T+E	M+C
VETER	O	T+E	M

AN INVESTIGATION INTO THE PERFORMANCE OF FIRST YEAR STUDENTS IN MATHEMATICS AND STATISTICS AT A SOUTH AFRICAN UNIVERSITY

Karl Stielau
Department of Statistics and Biometry, University of Natal
P.O. Box 375, Pietermaritzburg 3200, South Africa

1. Introduction

Statistics departments, and indeed universities, in South Africa are at present in a considerable state of flux. If one examines the state of tertiary education in South Africa (S.A. henceforth) ten years ago the changes are evident to even a casual observer. Although for many years the traditionally liberal campuses have defied apartheid legislation in formulating admission policies, it is a fairly recent phenomenon that universities such as Natal, Cape Town, Rhodes and Witwatersrand have clearly stated their intention (often as a Mission Statement) to make both the student body and the staff correspond as closely as possible to the community in which each university finds itself. This has resulted in a long overdue change in the type of applicants for entrance, and in particular to Statistics courses, that these bodies receive. In Natal the position is exacerbated by the large number of school education departments that issue matriculation certificates to school leavers; these are meant to allow the individual to enter a university without further examination. The proliferation of such departments is a glaring legacy of apartheid and their incorporation into a single education department appears certain under a new political dispensation. During 1991 and 1992 the Statistics & Biometry Dept. at the University of Natal admitted students from: Natal Education Dept.; Department of Education and Training; KwaZulu Education Dept.; Joint Matriculation Board; House of Representatives; House of Delegates; Cape,Transvaal and Orange Free State Education Dept.'s; as well as a large number of "homeland" departments, such as Ciskei, Transkei, Venda, Bophutatswana and adjoining states, in particular Zimbabwe.

What are the consequences of these facts to admission policies? Several South African universities have recently adopted a system whereby affirmative action is implemented. White students who have enjoyed a privileged and trouble free education, along with the so-called Indian and

Coloured students, are treated on the basis of points achieved. Black students, who have often experienced disruptions to their schooling such as burnings, loss of teachers, disruption of home life, are assessed on the basis of "potential to succeed" as evaluated by experts. Nevertheless, the "points system" is often used by faculties and departments to make recommendations to students on which courses to pursue.

2. The "points system"

Students who write matriculation examinations under any education body are awarded points according to which mark was achieved at a particular grade. Scoring is as follows:

	A	B	C	D	E	F
Higher grade:	8	7	6	5	4	3
Standard grade:	6	5	4	3	2	1

Marks below an F do not contribute to a student's score.

A decision on the suitability of a student for a particular course is then commonly based on points achieved, either total (for six subjects), Mathematics or English. It should be stressed that although black students are not formally subjected to a points evaluation for entry into university the points that such a student achieves in a subject are nonetheless informally considered in deciding whether or not a student should be admitted to a given course in Mathematics or Statistics. This practice is subjected to scrutiny in this paper.

3. The sample

Students considered here were registered during 1991 and 1992 at the University of Natal for courses in Mathematics (QM110) and Statistics (QM120). Although some students did one course only, attention is restricted to those who did both. Student numbers were roughly 500 in each course. Two outliers were rejected from the sample: the first was a student who had been without a teacher for over a year after the closure of his school; the second was a mature student who had matriculated in 1963 under quite different conditions. It could be argued that these anomalies occur in the correct proportions and should thus be included in the analysis. It is not thought that this would greatly affect the conclusions below. For the sake of brevity only the 1992 data are analysed below, but results were very similar for 1991.

4. A look at the data

During 1992 students registered for the courses concerned belonged to four faculties: Arts, Social Science, Agriculture and Commerce. They can thus be considered as representing a fair cross section of the student community, although due to degree requirements the majority were Commerce students.

It was felt relevant to consider the following variables: Matric marks for both Mathematics and English (as language competence was thought to be a possible factor in a student's understanding of lecture material), QM110 mark at university, QM120 mark and a mean of QM110 and QM120 reflecting general success in Mathematical/Statistical subjects. This mark is called QM henceforth.

Let us examine as a preliminary step in the analysis the distribution of marks achieved for QM110 and QM120 and the QM mark indicative of overall performance. These are shown below as Figures 1 and 2.

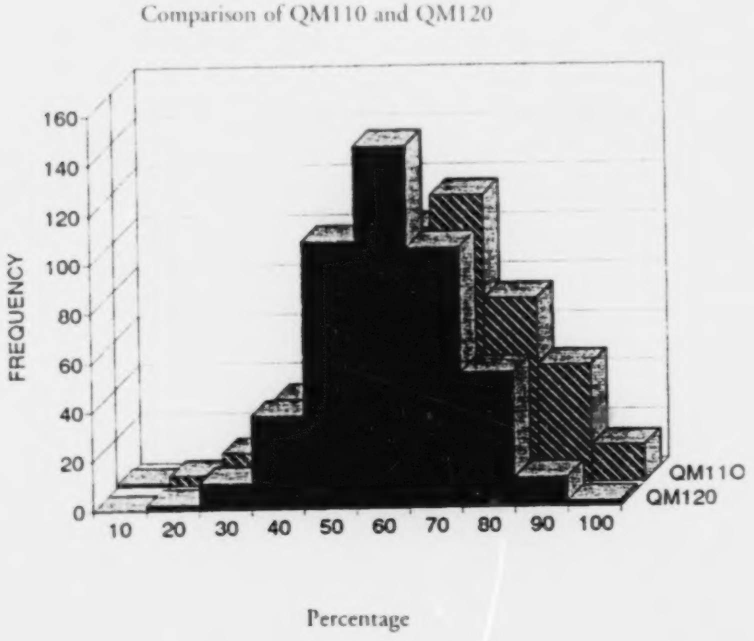


Figure 1. *Histogram of QM marks*

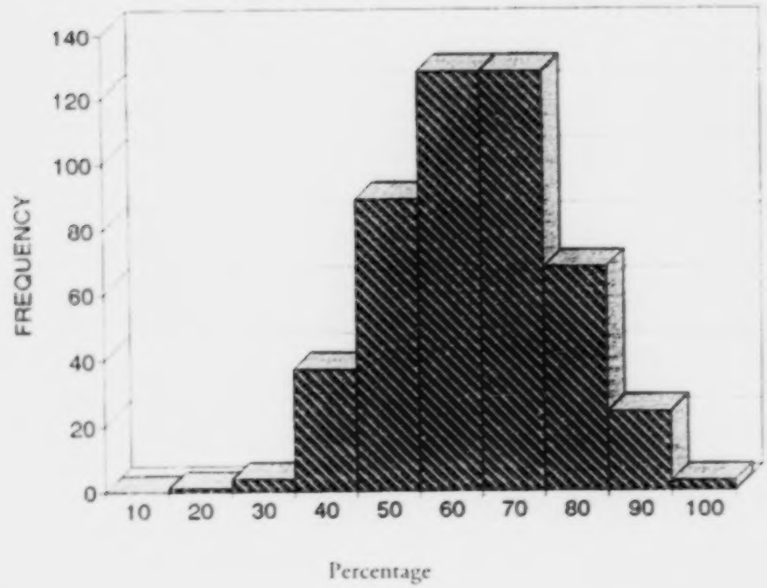


Figure 2

There is a slight skewness evident in the distributions; also it is evident that students fared somewhat less well in Statistics (QM120) than they did in Mathematics (QM110). The sample size was 415 after excluding students who for one reason or another were not doing both courses. Summary statistics are shown below:

	QM110	QM120	QM
Mean:	61.66	56.48	59.07
Median:	62	56	59

	MATRIC MATHS	MATRIC ENGLISH
Mean:	5.49	5.89
Median:	5	6

Correlation of QM110 with QM120 = 0.68 ($P<0.001$)
Correlation of QM with Matric Maths = 0.62 ($P<0.001$). Correlation of QM with Matric English = 0.21 ($0.01<P<0.05$) (all on 413 d.f.).

It should be noted that the pass mark is officially 50% although students scoring 49% are inevitably passed.

We now turn to the usefulness or otherwise of the points scored by a student in Maths or English at Matric level in predicting the mark of such a student in university courses in Maths and Statistics.

5. The models

Some of the analysis below constituted part of a scheme in which promising students at the University of Natal are financially rewarded for involving themselves in the solution of statistical consultation work related to their syllabi. Such work is carried out during university vacations and is supervised by a member of staff. In this paper the analysis is thus restricted to fitting straightforward linear models of the type where the expected QM mark is expressed as a linear function of one or more independent variables with unknown parameters.

6. Data analysis

As an initial step it was decided to keep things simple and thus to focus only on Matric points for the two subjects as independent variables, therefore ignoring grade (Higher or Standard) and all interactions, as well as education authority under which a student matriculated. The preliminary analysis is summarised below:

```
a) Only Matric Maths as X          b)  Only Matric English as X
   CONSTANT      22.9266               CONSTANT      41.2692
   SE OF Y EST   10.4252               SE OF Y EST   12.9370
   R SQUARED      0.3799               R SQUARED      0.0450
   X COEFF        6.5872               X COEFF        3.0186
   SE OF COEFF    0.4141               SE OF COEFF    0.6840
                     c)  Maths as X1 and English as X2
                         CONSTANT         18.2788
                         SE OF Y EST      10.4025
                         R SQUARED         0.3841
                         X COEFF'S         6.4143        0.9487
                         SE OF COEFF'S     0.4260        0.5669
```

It is a little surprising that Matric English appears to be of so little value in predicting university marks in the areas of Mathematics and

Statistics. The Matric Maths mark seems to perform in a similar fashion to the results reported by Fresen (1990).

It was now thought desirable to extend the analysis to incorporate other factors into the analysis. These included in particular the grade (H or S) on which each student wrote his/her Matric examination as a factor in the analysis. A variety of models was also fitted incorporating interactions. The highest R squared value achieved was for the model incorporating points for both subjects, grades for both subjects and interactions; it is however interesting to note that this model achieved an R squared value of only 0.418, thus very little better than the simple one based on Matric Maths alone.

Multivariate analysis by principal components in order to examine the effect of "department" under which a student matriculated was not highly successful owing to the large sample size. Research on this topic continues.

An attempt to obtain the conditional distributions university marks for QM (on Matric points achieved) was carried out. The results closely correspond to those of Fresen (1992) at the University of the Western Cape and are not reproduced here for reasons of space. Details and data sets are available on request to the author.

7. Some Conclusions

The results obtained in this study for 1992 were very similar to those for 1991, and are extraordinarily similar to those obtained at the University of Cape Town, the University of the Western Cape and the University of the Witwatersrand (Fresen, 1992). There is reason to recommend that Matric marks and points should be treated with considerable scepticism in the context of South African education at present. The predictive value of such marks/points for two fundamental subjects (English and Mathematics) is clearly low. It is to be hoped that recommendations for course or faculty entrance will not in future be based on such predictors; this will become more and more important as South Africa's turbulent education systems coalesce and develop, along with the creation of a new political dispensation.

Acknowledgements

K. Stielau wishes to thank the University of Natal's Internship Programme for sponsoring the invaluable work of Ms Reshina Roy during 1992. The author also wishes to thank Mr J.B. Levin for assistance with computer programming.

Bibliography

Fresen J. and Fresen J. (1990), *Technical Report*, University of the Western Cape.

Fresen J. (1992), *Unpublished conference paper*, S.A. Statistical Association Annual Conference, Port Elizabeth.

TEACHING BIOSTATISTICS TO MEDICAL STUDENTS

Roel van Strik
Institute of Epidemiology and Biostatistics
Erasmus University Medical School
PO Box 1738, 3000 DR Rotterdam, The Netherlands

Statistical reasoning is applicable to all areas of medicine. The following few examples may serve to illustrate this.

Example 1. Reference ranges of "normal values" for various constituents in human plasma are to be based on accurate measurements in sufficiently large random samples of healthy individuals from relevant populations. Such a range is usually estimated from suitable percentiles of the observed frequency distribution, either non-parametric or based on a specific model like the gaussian distribution. The resulting ranges may be used by medical practitioners as an expedient to establish a diagnosis for an individual patient and later on for monitoring the course of the disease during a particular treatment regime.

Example 2. Evaluation of the merits of a new diagnostic test procedure either to confirm or to exclude a suspected disease, should be based on random samples from two populations, the first consisting of individuals with the disease truly present according to a so-called 'gold standard' and a second population of individuals with the disease truly absent by the same standard. Considering a certain diagnosis, a positive test should indicate that this is likely and a negative test that this is unlikely with respect to a particular patient. The merits of such a binary test are expressed in terms of sensitivity, specificity and predictive values, which may all be interpreted as conditional probabilities. In particular a predictive value may be derived from the values of sensitivity and specificity of a test together with the prevalence of the disease in the population concerned, by using Bayes' theorem:

$$P(D{+}|T{+}) = \frac{P(T{+}|D{+})*P(D{+})}{P(T{+}|D{+})*P(D{+})+P(T{+}|D{-})*P(D{-})}$$

or

$$pred.\ value = \frac{sens*prev}{sens*prev+(1-spec)*(1-prev)}$$

Example 3. Critical appraisal of the efficacy/safety balance of a new medical treatment requires, of course, relevant and reliable data. As a consequence careful design, execution and analysis of such a medical research project is required. The importance of this is clearly recognized by drug regulation authorities in Europe, Japan and U.S.A., as witnessed in various guidelines, for example the one by the European Commission on Good Clinical Practice, which states in its foreword: "good statistical design is an essential prerequisite for credibility of data". The best appreciated experimental design in clinical research, is the "parallel groups randomized controlled trial", particularly when definite answers are looked for. If a two-period cross-over comparative trial design is at all feasible, this comes next, but usually at a more preliminary stage of a research programme.

These three examples may suffice to explain why statistical methods are used abundantly in research papers in the medical literature. A medical doctor who wants to keep abreast of scientific progress in his field in order to strive for optimal patient care, will therefore be confronted with various aspects of statistical reasoning in his professional literature. Whether or not he has received any formal education in medical statistics he will anyway be involved in statistics, either only as a consumer or also as a producer when he is actively doing medical research and wants to publish his results. In the latter case, this may sometimes lead to peculiar applications of statistical models as illustrated with the following two examples.

Example 4. Some years ago, a paper in an international journal on growth and reproduction contained the graph shown in Fig. 1, which may be considered as an obvious example of statistical mis-modelling. It shows a linear regression model fitted to apparently heterogeneous data from 31 females. To point out albeit superfluously, some of the peculiarities:
- six distinct subgroups are involved, clearly different as to mean, variability and slope, but nevertheless pooled when fitting this model;
- values of $y - x$ are plotted against x, implying a built - in slope which is, of course, negative if x and y are uncorrelated;
- according to the text, for one subgroup (closed triangles) individuals with x - values below 60 were excluded;
- a straight line was drawn even though both axes show interruptions either with or without change of scale.

Example 5. A paper in a medical journal contained a table showing data of a double blind efficacy study with a drug against insomnia (see

Table 1), which is clearly a specimen of sample size inflation. It bears on the results obtained during a seven days study with 78 patients. Obviously the study unit was not recognized since the seven answers from one patient were clearly counted as if given by seven different individuals. Moreover about 14% of the data were apparently missing, viz. 18% on drug and 11% on placebo.

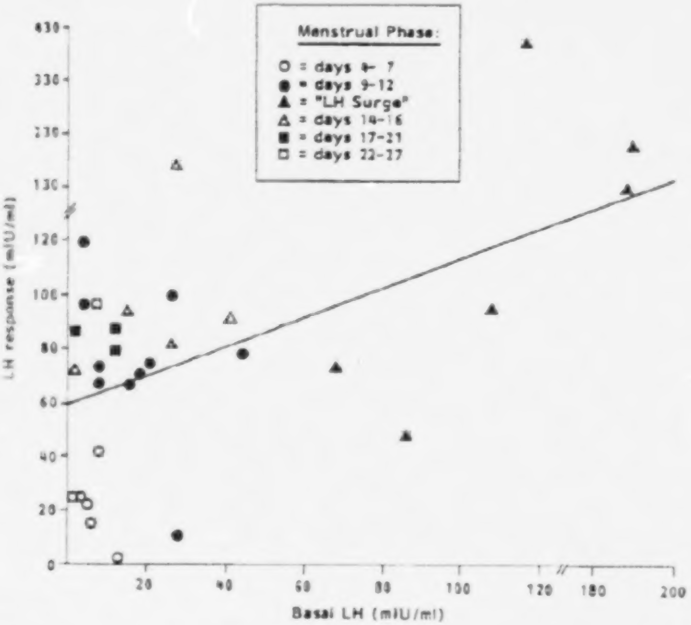


Figure 1. *Scatter diagram with regression line showing positive correlation between basal LH levels and response to LH-RH ($r = 0.55$, $p < 0.01$, $y = 58.8 + 0.8\,x$) throughout the menstrual cycle.*

Table 1. *Quality of sleep: Did the medication help you to sleep?*

	Helped	No help	Total
Drug (n = 40)	216	13	229
Placebo (n = 38)	150	92	242
Total	366	105	471

$\chi^2 = 70.84$; $P < 0.001$

These two examples may suffice to indicate that a basic knowledge of biostatistics is of prime relevance for every medical doctor. This brings us to the questions: when, what and how to teach biostatistics to medical students?

When exploring a way to successful teaching of biostatistics we may start by considering two theses:

- Medical students are generally well motivated but mainly for the study of medicine and not so much of chemistry, physics, statistics, etc.
- In order to practise medicine effectively in their later career medical students will need to know enough about biochemistry, biophysics, biostatistics, etc.

These statements may raise some questions:

- How to motivate medical students to take the study of biostatistics seriously?
- How to tailor a biostatistics curriculum to attain coverage of relevant subjects, make it interesting for medical students, and ultimately worthwhile in their opinion?

A successful method of approach to cover the second question, will most likely contribute to a favourable outcome with respect to the first one.

Let us now take up the question: When to teach? After what has been discussed so far, it may be argued that teaching biostatistics should be dispersed throughout the whole medical curriculum. However, the practice of everyday life in a medical school will in general set fairly strict limits to that idea. Moreover in our Medical School at the Erasmus University in Rotterdam, like in most medical schools, the medical curriculum contains also the related subjects epidemiology, social medicine, public health, clinical decision analysis and medical informatics. These subjects are taught partly parallel and in association with biostatistics and partly in other years. In this context we attained a satisfactory compromise: we are teaching obligatory courses in first year and in second year classes, and elective courses in later years.

The next question is: What to teach? In our opinion, given three different periods of study, teaching can be tailored to three levels of understanding:

- Initially the students learn to understand the line of thought in the application of statistical principles to the design and analysis of medical research; this part emphasizes a.o. study designs (including definition of study unit), data description, probability laws (applied to diagnostic problems), sampling distributions, statistical inference (confidence intervals; hypothesis testing; parametric and non-parametric; one and two samples); simple regression and correlation.

– In the second period the students read medical research papers and make a critical appraisal with respect to methodological issues; this part includes a.o. additional aspects of study designs, analysis of variance, evaluation of measurement methods, multiple and logistic regression, survival analysis.

– At the third stage the students, usually being involved in a (smaller part of a) research project, should practise the judicious application of methodological principles regarding design, analysis and interpretation; at this stage emphasis may be given to a special subject e.g. randomized controlled trials as a research design model in medicine.

With respect to the question: How to teach? I like to submit for your consideration a few general guidelines.

– Teaching biostatistics to medical students requires a well-balanced combination of lectures, tutorials and practicals. From our experience we strive for about an even distribution of them.

– When lecturing the teacher should resist the temptation to discuss statistical techniques as such, illustrating their use on invented data with a statement like "these are assumed to be realisations of indepedently distributed random variables".

– It is advisable to use actual data and to elaborate on the origin of the data and the way they were collected, stressing that this should flow from the medical question and from the design of the study. The student should thereby learn to formulate a plausible albeit tentative model of the data in concordance with the actual data generating process, as a basis for the statistical methods to be applied in order to (hopefully) answer the medical question in a reliable way. In our teaching we use as an initial and fairly simple example for this, a video film showing a doctor measuring systolic and diastolic blood pressures on twelve occasions. However, the students just hear the heartbeats as if from the stethoscope and see simultaneously only the scale of the manometer in the film, from which they have to read both blood pressure values. Their readings, as recorded on OCR-forms by the students and then further processed, form a unique data base. The twelve measurements, for example, of diastolic blood pressure represent six patients each measured twice in randomized order and blinded for the students. This makes it feasible during practicals to have students answer questions about inter-observer as well as intra-observer variation, while sitting at their PC's and using standard statistical software. They can easily produce histograms and descriptive statistics, test hypotheses, construct confidence intervals, perform regression and correlation analyses, etc. The students are usually well motivated to this, since their own data are involved and they know exactly how these were assembled.

BEST COPY AVAILABLE

– With respect to statistical inference, the student should be convinced of the relevance of confidence interval estimation of clinically important effects, instead of merely reporting P-values.

In conclusion, teaching of biostatistics to medical students should be carried out in a way:

– that makes clear that biostatistics is more than a "bag of tools" for data analysis per se;

– that does not leave the impression with the students that only the shape of the data matrix is important;

– that will overcome their lack of interest and motivation and instead convince the majority of the relevance of statistical methodology to medical practice and research.

EDUCATION IN STATISTICS AT THE HIGHER INSTITUTE OF CHEMICAL TECHNOLOGY - SOFIA

Kamen Velev, Iunchev Vuchkov and Hristos Iontchev
Department of Automation
Higher Institute of Chemical Technology
Bul. Kl. Ochridski No 8,1756 Sofia, Bulgaria

1. Organisation of the education in statistics at the Higher Institute of Chemical Technology - Sofia

The Higher Institute of Chemical Technology-Sofia (HICT-Sofia) is concerned with undergraduate and postgraduate education in chemical technologies, metallurgy, material sciences and automation. The teaching of statistical methods for engineers began more than 20 years ago and HICT-Sofia is widely recognised as a pioneer in disseminating the statistical way of thinking among the engineers of Bulgaria.

Education in statistics at the HICT-Sofia has several levels depending on the specialisation chosen by the students. All students (except those specialising in automation) obtain some basic knowledge in probability theory from the courses of mathematics in the first two years of study. Then they have some deeper insight in applied statistics, following the course of "Theory of experimental research and optimisation".

The main objective of this course is to give the students an introduction to model building, regression analysis, design of experiments, hypothesis testing and optimisation methods deep enough to make possible their application in solving technological problems. Our experience showed that the course is extremely useful for students in chemical engineering, metallurgy and material sciences but it requires more hardware and software support now.

Students in automation obtain more advanced knowledge in statistical methods and their application. They follow several courses with statistical content like "Statistics for engineers", "Theory of experimental research", "Identification methods", "Quality control" and some others where partially statistical methods are applied.

Some of these courses are supported by problem oriented software developed by the academic staff of the Department of Automation.

A considerable part of diploma thesis and postgraduate research is often connected with application of statistical methods for solving technological problems.

To satisfy the need of teaching statistics for postgraduate students and industrial practitioners a special curriculum for statistical education was included in the programme of the School for Postgraduate Education in Cybernetics at the Department of Automation. For more than 15 years it has trained a large number of postgraduate students and practitioners from all branches of industry in applied statistics and quality control. In the last few years the crisis in industrial production led to nearly closing this way of teaching.

The education in the regular courses and the postgraduate education was supported by a variety of special textbooks and manuals in statistics and design of experiments written by the departmental staff. A Catalogue of Optimal Experimental Designs was also developed and published. It is still widely used in the experimental practice and for education (Vuchkov *et al.*, 1978).

2. Prospects in the education in statistics in HICT

The low quality of industrial production in Bulgaria is one of the main restraints to the process of transition to a market economy. A substantial contribution to solving this problem can be achieved by the wide application of statistical methods for quality improvement such as industrial design of experiments and optimisation methods. In a short time substantial numbers of students and industrial practitioners have to be trained to a competent level in this area. This leads to the need for highly qualified university teachers who are familiar with the successes of the leading European countries.

The development of education in this area has been held back in the last few years by the need for sophisticated computer techniques, software and teaching materials.

One of the possible ways of improving the education in statistics is European cooperation within TEMPUS. In the last two years the staff of the Department of Automation has been involved in two TEMPUS projects. The cooperation will promote professional contact, the transfer of expertise, sophisticated hardware and software and also technology transfer between United Kingdom, Italy, Germany and Bulgaria. The major aim is to reconstruct university education in the subject area in Bulgaria to bring it into line with European standards by transferring educational expertise and know-how from EC countries to Bulgaria.

It concerns the organisation of educational units and the management and conduct of undergraduate and postgraduate education. Taking into account that the university system of Bulgaria is changing now in

BEST COPY AVAILABLE

accordance with the transition to a market economy it is a recognition of an urgent existing need in Bulgaria. Academic staff retrained in the EC countries immediately apply their knowledge to develop teaching materials for wide distribution in the country. They will also teach large groups of undergraduate and postgraduate students and industrial practitioners. This will have a long term multiplier effect on education and on industrial applications and will contribute to the economic restructuring of the country.

The TEMPUS project connected mainly with statistical education has a title "Education in industrial design of experiments and optimisation of production in Bulgaria" (Torsney *et al.*, 1992). Its main objective is to develop in the HICT-Sofia an educational centre for teaching undergraduate and postgraduate students and practitioners from all branches of industry. The following outputs are to be achieved:

– Improving the level of undergraduate and postgraduate education in the area of application of statistical computing, experimental design and optimisation methods in solving industrial problems and promoting the transition to a market economy.
– Creation of new up-to-date facilities for education and production of teaching materials with a long lasting multiplier effect.
– Provision of retrained university teachers who in turn will promote the retraining and education of other teachers, of postgraduate students and of industrial practitioners.
– By improving the educational level in statistics in technological universities in Bulgaria to increase the speed of social and economic changes.
– Bringing the level of education in Bulgaria nearer to that of the EC countries and continuing the cooperation between the universities in EC and Bulgaria.

Some of the JEP activities in Bulgaria are the following:

– Two industrial courses on "Strategies in Computerised Education" and "Application of Industrial Design of Experiments and Optimisation Techniques for Improving the Quality of Production".
– The update of three existing training courses in "Methods for Experimental Research and Optimisation" for undergraduate students in HICT-Sofia. Retrained Bulgarian Staff will take an active part in developing new teaching materials, in updating the curricula, in exchanging information and in introducing a European dimension. A Catalogue of Optimal Composite Sequential Designs is developed as teaching material.

- A laboratory provided with 10 personal computers connected in network is developed and will be intensively used in educational activities. Computers and software are compatible with those presently existing in UK centres to maximise technology and know-how transfer.
- An integrated Postgraduate Training Unit is developed. It consists of a seminar room and office facilities equipped with a computer, with desk-top publishing and graphics software and also copying and printing facilities to enable efficient preparation of educational manuscripts and slides.

The establishment of the Laboratory and the Postgraduate Training Unit will enable the expertise and curricula of EC countries to be transferred to Bulgaria in a short time. It will result in developing the Higher Institute of Chemical Technology -Sofia into a national centre for the education and distribution of expertise in "Design of Experiments", "Statistical Quality Control" and "Optimisation of Industrial Production" in Bulgaria.

Bibliography

Torsney B., Iontchev H. and Giovagnoli A., (1992), *TEMPUS JEP- 3893-92. Education in Industrial Design of Experiment and Optimisation of Production in Bulgaria*, a report.

Vuchkov I., Iontchev H., Damgaliev D., Tsotchev V. and Dikova D. (1978), *Catalogue of Sequentially Generated Designs*, MNP Publishing House, Sofia.

COMPUTERS, VIDEO AND OTHER TOOLS IN THE TEACHING OF PROBABILITY AND STATISTICS

Invited papers

Tim Arnold
Rolf Biehler
Andrej Blejec
Clifford Konold
David S. Moore

Contributed papers

Furio Camillo *et al.*
Claudio Capiluppi *et al.*
Steve Cohen *et al.*
Joe Griffin *et al.*
Michel Janvier *et al.*
Peter Jones *et al.*
Jiří Militký *et al.*
Annie M. Morin
Juha Puranen

THE ROLE OF ELECTRONIC COMMUNICATION IN STATISTICS EDUCATION

Tim Arnold
Department of Statistics, North Carolina State University
Box 8203 Raleigh, NC 27695, USA

1. Current problems in statistics education

Many researchers and writers have pointed out that there is a great need to improve methods of statistics education (Bisgaard, 1991; Cobb, 1991; Snee, 1993). Statistical literacy has become a social imperative if we are to share a responsible, thoughtful citizenry. However, often in the classroom students are isolated, working in silent interaction with printed material, or working through "cookbook" computer homework assignments. It is as if in forcing the student to go through certain motions, learning will somehow magically happen. But, as has been pointed out previously (Riel and Levin, 1990), students must be active learners, engaging themselves in the exploration of ideas and communicating those ideas to each other and, at the same time, to themselves. The value of active participation and student collaboration is well known; the problem is in how we structure a learning environment to gain that value.

Today's students also need skills of expression, skills for sharing ideas and information. Seldom does the statistician need only an answer and not need to communicate that answer. It seems apparent that as statistics educators, whether of statisticians or statistically literate citizens, we must be concerned with teaching methods for both creating and communicating statistics information. In foster these communication skills, students must be allowed to practice communicating in the classroom. Indeed, educators are becoming more aware of the importance of social interaction in the classroom (Riel, 1990). There are problems in making these changes, of course; teachers cannot reorganize their classes to meet these goals and still maintain the traditional classroom form. Additionally, current world-wide economic pressures leave teachers with tight budgets, the need to do more with less, often resulting in larger classes and even more difficult communication problems.

Technological advances now make it possible for teachers to consider the possibility and promise of electronic communication through computer networks. This technology can enable teachers to form new relationships with their students and with other educators around the world, providing much-needed mentoring and support. Networking can also provide an outreach function for nontraditional students, dissolving barriers of geographic location and at least some cultural impediments. Electronic communication allows students to offer each other encouragement, both academic and personal, often resulting in a more positive evaluation of the classroom experience (McCormick and McCormick, 1992).

2. Current forms of electronic communication

Electronic communication today can take several forms, from long distance telephone service to video-conferencing, fax services to slow-scan (freeze-frame) television. All of these systems are designed to deliver data, voice, text, or image from one point to another. While these forms of electronic communication are powerful and hold much promise for educators, the focus here will remain on the most widespread and potent forms for the classroom: electronic mail, computer conferencing and methods of globally-networked "computer-mediated communication".

Many universities, companies, colleges and private individuals are linked together by various international computer networks collectively called the Internet. While it is difficult to estimate the actual numbers comprising this network due to its rapid growth, it is safe to say the Internet is composed of over 5,000 other networks in over 107 countries (Roupp and Pfister, 1993). It is growing from 10 to 15% monthly. The Internet began as ARPANET (Advanced Research Projects Agency Network) in 1969, developed to help university researchers share information. Since then it has grown to include such large networks as NSFNET and NEARNET. In addition, many networks such as BITNET are tied to Internet but are not integral parts of it. An individual user with an account on a computer linked to the Internet gains access to resources not only on that computer, but on any Internet computer that offers public information. There are many such sources; some users have likened the task of getting information on this knowledge-rich network to "drinking from a fire hose" (Gelernter, 1992). The user also has electronic-mail access to all other Internet users, and to users on networks that have connections to the Internet.

E-mail. One of the most useful tools for obtaining information from the network is electronic mail (e-mail). E-mail is essentially a one-to-one or one-to-many form of communication. Besides providing person-to-person messaging, e-mail is used in electronic discussion groups. In these groups, any message sent to a specific central address is automatically rebroadcast to all members of the mailing list group. These discussions bring people together as a focus group, enabling them to consider common problems, share solutions, and argue issues. EdStat-L is one such forum, dedicated to providing a vehicle through which techniques, tools, ideas and philosophies of statistics education can be shared.

E-journal. Network-based electronic serials are periodicals that are primarily "published" on computer networks like the Internet. Several of these function as electronic journals (e-journals). A scholarly e-journal is quite similar to a paper-based journal, containing refereed articles, reviews, and other information. However, the two differ in several ways. The most obvious difference is that an e-journal exists on the network instead of on paper. Readers access journal articles and other information by e-mail request or direct access to the archiving computer. Other differences include the ability of an e-journal to offer databases or other information associated with an article, as well as the ability to interact with and conduct searches on the entire journal archives.

Computer conferencing/Bulletin Board Systems (BBS). In contrast to e-mail, computer bulletin boards and conferencing are methods for sending messages to an area or message database on the network that is shared by others at various locations on the network. All users must access the same network. Some conferencing software is sophisticated, allowing users to search, branch comments into topics and replies, and use a number of other functions. Users generally post information that is later read in sequence by other users accessing the conference. UseNet News is one type of conferencing system used around the world. It is extremely popular, with users debating on over 2,000 subjects, called "newsgroups". The public-domain software also allows local newsgroups or conferences to be set up for classroom or departmental communications.

Methods of direct access. Gopher and World-Wide-Web are two software programs available at no charge that allow the user to use his computer as an information "client", requesting information from any of the several thousand information "servers" on the Internet. Because there are so many public access information servers, many of which are devoted to specific subjects, these programs are needed to help navigate the sea of information available on the network. The programs can employ interfaces that use hypertext, graphics images, sound files, and text. For connecting to public information servers these programs are the current state of the art.

3. Human factors in computer-mediated communication

Regardless of the specific technology employed, computers enable a new form of mediated communication with psychosocial properties that are quite different from those associated with other forms of communication. The computer's impact on communication extends well beyond the convenience of faster and more efficient information transmission (Harrison and Stephen, 1992). Because the system allows the user to delay both receiving and responding to the message to suit his/her schedule, he/she controls the time of communication.

Computer-mediated communication enables collaboration between learners by removing time and distance barriers. Students can help each other in an active collaborative environment and instructors can provide further opportunities for interaction with on-line seminars, group discussions or joint experiments. This type of communication allows for interaction both between students and instructor and among the students themselves that would be impossible in a conventional course.

This mode of communication is also inclusive in that it can incorporate many voices and points of view. This affords every student the opportunity to have an equal chance to participate, independent of status and role differences. The traditional paradigm of lecturing professor/silent student is relaxed; there is no explicit "downward" communication channel. The inherent structure of the network is horizontal (Lewis and Hedegaard, 1993; Roupp, 1993). Conversations may be many-sided and can contain multiple threads of content. These information exchanges are intertwined with the learning and teaching activities that provide much of the subject matter for discussion (Drayton, 1993; Kaye 1992). Conversations are automatically archived so they can later be scanned or referenced. When the use of network communication is frequent, the system provides a real forum, a round-table discussion with many points of view. This gives us the ability to reach beyond the traditional classroom boundaries and, hence, can change both the size and composition of the educational community.

4. Examples and resources: integrating electronic communication and statistics education

Distance education. Electronic communication can be used to supplement traditional courses or provide the primary vehicle for distance education courses. Students engage each other and their instructors in online discussion groups, reading, thinking, then responding. This structure encourages a much more active learning experience than traditional classroom methods. Distance learning is particularly useful for the disabled student, allowing interaction that might be difficult if not impossible otherwise. Bellman (1992) points out that this type of classroom has few prejudicial elements, and so enables more assertive involvement in the education process by traditionally apprehensive or passive learners.

There are several successful examples of electronic communication and distance learning today. The Appalachian Community Service Network (ACSN) combines broadcasting, satellite, cable television and telephone technologies to provide instructional programming to Appalachia, a rural part of the United States that faces some of the same problems as many emerging nations (Van Deerlin, 1980). The Global Schoolhouse Project, sponsored by the U.S. National Science Foundation, uses live two-way interactive video on the Internet to reach classrooms in California, Virginia, Tennessee and the United Kingdom. This allows students to extend their learning experiences far beyond the traditional classroom by allowing them to interact and collaborate with their peers regardless of geographic location. In Montana, the Big Sky Telegraph Project uses a combination of network technologies to reach over 116 one-room rural schools. The system uses packet-radio broadcasting (wireless network communication) to transmit audio, text, and animated color graphics over a 50 mile range (Utsumi and Villarroel, 1992).

There are also dangers to consider in distance education projects, however. The social changes that this technology brings into our lives and classrooms must be carefully considered. Fundamental changes in human behavior, regardless of the setting or the goal, are difficult to achieve. For distance education to work, much more may be needed than simply technical infrastructure, especially when the distance is between cultures. In online communication, since words are the only information available, the opportunity for multicultural miscommunication is ever present. Shades of meaning and style of communication become even more important since visual and auditory cues are absent (Shapard, 1990). Debate may seem like an attack to some participants, or advice may sound like arrogance. Although students and instructors may not be experts in the social interpretations of online communications, it is safe to say that human decency is always in style. However, online communications may require more of it, or more thought about how not to violate it, than other forms of communications that contain more nonverbal information.

Collaborative learning opportunity. The structure of computer-networked communication allows great opportunity for collaborative learning. The network may be used to perform online experiments, perhaps linking geographically-distant classrooms for joint experiments. The multitude of information already on the Internet can be used in a "data hunt" in which student groups access the network to find datasets related to their interests, retrieve and analyze the data, reporting back results (or problems) via the network. Other groups on the network may be surveyed by a class, or called upon for information by student groups. "Experts" or "clients" may visit the class online for an "electronic consultation", explaining an industry application of methodology, or describing a problem. Such visits might even be scheduled live, so students could interactively ask questions or posit solutions with the visitor.

In exploring and experimenting with networked collaborative learning, it is important to keep in mind that the philosophy of "if you build it, they will come" is a recipe for failure. One of the most important factors for success for computer conferencing in the classroom is a commitment to leadership that supports the activity (Canning and Swift, 1992). Indeed, Reil and Levin (1990) point out that the goal must be shared by the group, access must be easy and equally available to all and someone must act as the facilitator or leader. If the students fail to see the practicality of using the network, or if they are not explicitly encouraged to use it, they will not use it.

Nuts and bolts. Perhaps more mundane, the use of the network in improving the classroom infrastructure, such as student feedback or test and quiz results, can be quite helpful to both student and instructor. In the Statistics Instructional Computing Laboratory (SICL) at North Carolina State University, a "public" area is set aside on the network for instructors to put electronic copies of exams or quizzes from previous years, as well as results of exams and quizzes given during the semester. Other materials may appear there, such as special datasets for the students to experiment with, or pointers to other information on the network. Some instructors encourage group interaction via classroom bulletin boards, posting homework assignment errata or helpful examples similar to homework assignments. Other instructors use the "ghost" login facility, whereby any student logging in under the name "ghost" can send an anonymous message to the instructor. The intent of this facility is to encourage honest feedback, suggestions, and commentary on the course. This is a popular feature that has resulted in instructors receiving some thoughtful criticism and suggestions. Finally, all locally-written help manuals for the computer system and software packages in SICL are available through the lab's online help system. All of this material, from computer-command descriptions to notes on multiple regression, may be searched by keyword or accessed sequentially via the menu system.

Mentoring and electronic community. One of the most important facets of electronic communication is the ability to create communities, not by a common geographic location, but by a commonality of ideas. The electronic discussion groups on the Internet are well suited to this task. Statistics educators may find a congenial community in EdStat-L, the discussion group for statistics education. The purpose of the group is to provide a forum for comments, techniques and philosophies of teaching statistics. The group's stated goal is to bring together every teacher, student, researcher, and specialist interested in improving statistical instruction. The group's communications range widely, sometimes debating the issue of what academic groups are the proper teachers of statistics, sometimes discussing different ways of looking at (and explaining) certain statistical concepts, and sometimes just "chatting" with each other, all in a highly interactive, informal coffee-house style. This facility can also be used as a way for experienced teachers (mentors) to pass down advice, teaching anecdotes and the wisdom of experience to younger, beginning teachers. For many isolated statisticians and statistics educators, the network can be a place to find a much-needed community.

More formal interaction with the international statistics education community is also available on the network: The Journal of Statistics Education (JSE) is a new, rigorously-refereed electronic journal on post-secondary statistics education. JSE publishes high-quality articles on a variety of topics related to the teaching of statistics. The electronic medium enables readers to download datasets and teaching materials, search interactively the entire archived journal, access articles immediately as they become available, or participate in the parallel electronic discussion forum JSE-Talk, a forum devoted exclusively to discussion of issues concerning the JSE readership.

The JSE Information Service, an adjunct to the JSE, is a growing electronic archive of information, software and discussions related to statistics and statistics education. In addition to the JSE itself, the Information Service includes the JSE Guidelines for Authors and Call for Papers, and fully indexed and searchable archives of several electronic discussion groups related to statistics or statistics education. Also available through the service are software related to statistics or statistics education, information and programs to help users find and use other network resources, the national archive of SAS macros and programs, and information about other services of interest elsewhere on the Internet.

5. Future possibilities in electronic communications

Though it is generally dangerous and unfruitful to try to predict the technological future, there do seem to be some developments in electronic communications that are inevitable.

The most obvious development to come is that of the network itself. More and more of us will become connected to the network, and hence connected to each other. The major problems in network communication today have to do with the incompatibilities existing between the current plethora of networks and software and hardware companies. The many networks, service companies and lack of standard communication protocols belie the fact that we are not in a communications revolution, but an evolution. The situation is similar to the growth of other products and services that had a large impact on society, such as telephone service, the automobile industry, or cigarette manufacture. These industries all went through natural stages of growth: early innovation and experimentation, proliferation of brands and product standards, and finally stable growth resulting in a small number of standards or product types, a relatively small number of companies, and rapid spread through the culture. In the beginning days of telephone service, it was sometimes necessary for subscribers to belong to different telephone companies-users had to be on the same network in order to communicate. This is not far from the situation today in computer networked communication, and it seems probable that the evolution will continue.

Videoconferencing is another technology that is still expanding into the education field, allowing teachers to reach students in geographically-dispersed locations. As technology has advanced, videoconferencing has moved from concept to executive privilege and is now becoming a practical communications tool. With further advances in technology this form of communication will no longer require large boardroom-sized facilities but will be usable from a desktop computer on the electronic network (Gill, 1993).

Virtual reality may become important to us as educators. Virtual reality has been described as "a three dimensional participatory multi-sensory computer-based environment that occurs in real time" (Randall, 1992). This technology makes possible in the fields of distance communication and education, something new called telepresence, a paradoxical term that might be defined as having or sensing one's presence at a remote or even virtual location. Techniques in education such as visualization and simulation could certainly be affected, given this power to create virtual learning environments. Virtual reality combined with the global network could give us the ability to form workgroups whose members are geographically-distant, but able to communicate within the virtual networked reality, a version of "cyberspace" (Reingold, 1991). Using this technology, students and teachers could physically interact with objective, graphical databases, perform rotations of graphic images by hand and work together in virtual simulations of experiments. Whether the technology will be able to transcend social barriers and deliver the many promises made for it now provides fuel for hot debate in today's popular press.

Another development that seems certain is that we will need to become even more information literate. This entails "the ability to realize that information is needed, and to locate, evaluate and use effectively the needed information" (Wegner, 1992). The sea of information available on the networks will affect the life of each of us, whether student, teacher, or citizen. In order to survive and prosper in the intellectual climate of the future we will have to rely on each other for training, assistance, and encouragement in the task of lifelong learning.

6. The prospect of a truly global network

Although the Internet is growing at a furious rate, and hundreds of thousands of users are now connected, the great majority of the traffic on the network comes from North America. There is no reason to doubt that other developed countries will soon be as thoroughly connected to the network. However, it is easy to forget that we are not the whole group because the only voices we can hear on the network are our own. The Lesser Developed Countries (LDCs) are sparsely connected to the Internet and they may encounter severe political, financial and social difficulties in achieving the connection. It is important to note that poor or remote locations, even in developed countries, will have similar difficulties in becoming connected to the network.

By learning the basics of how the technology can be used for advancing education, we will be better prepared to advise our colleagues or lead our classes to pioneer the future of networked education. The technological solutions are simple and relatively inexpensive. As a result of continuing improvement in computer and communication technology, a meaningful network can be built using low-cost appropriate technology. The two technical solutions described here are elegantly simple, inexpensive and widely used.

Grassroots networking by telephone: FidoNet. FidoNet is a wide-area network that uses modems on the direct-dial telephone network. There are over 13,000 nodes on this user-run network, making it even larger than the popular and well known BITNET network. While hobbyists predominate on FidoNet in North America, about half of the public FidoNet systems in Europe are run by small- to medium-scale businesses. In Africa, the network is heavily used by non-governmental organizations and poorly-funded academic institutions. Within North America, there is growing use within the primary and secondary school systems. FidoNet is structured so that any node can communicate with any other node without the aid or consent of technical or political groups at any level (Bush, 1992). FidoNet communication is limited to e-mail or broadcasts similar to electronic discussion groups. While messages may take days to reach their destination, the network makes it possible to communicate with some remote locations where no other form of communication exists. A computer, modem, the FidoNet software and access to a telephone line are the necessary components of the network.

Packet radio and low earth orbit satellites. The power of packet radio as used in the Montana's Big Sky Telegraph project to reach remote areas can be vastly increased when coupled with satellite technology. For example, VITA's (Volunteers in Technical Assistance) satellite or "orbiting mailbox" will cover every spot on the earth twice a day. The satellite orbits at an altitude of 800 km. When it passes over a ground station, the station uplinks or transmits messages to the satellite and the satellite similarly downlinks or transmits messages or files to the ground station. The solar powered ground stations, which are relatively cheap, then broadcast the digitized messages on a radio frequency to computers in the local area that are "tuned in" to the packet-radio broadcast (Ronkin, 1991).

7. Conclusion

As educators, we need to work to improve the techniques we use. We need to involve our students with their learning and foster better communication skills. We need to reach out to students who need education, but have been left out of the traditional classroom. The promise of electronic communication offers great opportunities for distance education, imaginative collaborative learning projects, and help in making information easily accessible. In addition to technology for classroom projects, there is a wealth of information and services on the network for statistics educators themselves, such as the EdStat-L discussion group and the electronic Journal of Statistics Education.

We must guard against letting technology change the underlying values of education, remembering the proverb, "When all you have is a hammer, everything looks like a nail". The overweening desire for technology can blind us to problems, perhaps making us forget our reasons for looking to technology in the first place, which is to improve our students' and our own learning. It takes more than technology alone to truly make a difference. Connecting everyone to the global network will have no positive effect by itself; our own imagination and effort is the real empowering force. However, it is important to know that technical solutions do exist, and that these solutions can have a large impact for little cost.

Armed with both knowledge and imagination, we can look to the future to take what is useful from the technological world in making the classroom a better place, engaging our students in their own learning. We can use the medium to reach out to others, and learn from others in ways that were before unimaginable.

Bibliography

Bellman B.L. (1992), Computer Communications and Learning, *New Directions for Teaching and Learning,* 51, pp. 55-63.

Bisgaard S. (1991), Teaching Statistics to Engineers, *The American Statistician,* 45, pp. 274-283.

Bush R. (1992), *FidoNet(tm)*: Use, Technology, and Tools, Network communication: Use anonymous FTP to dhvx20. csudh.edu, filename: global-net/fidonet

Canning C. and Swift K. (1992), Connecting the University and the Field of Practice: Computer Conferencing in Education at the University of Michigan, *Empowering Networks: Computer Conferencing in Education,* Educational Technology Publications Inc., Englewood Cliffs, New Jersey, pp. 1-33.

Cobb G. (1991), Teaching Statistics: More Data, Less Lecturing, *Amstat News,* 182, 1, 4.

Drayton B. (1993), Design for a Science Network, *LabNet: Toward a Community of Practice,* Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 145-197.

Gelernter D. (1992), *Mirror Worlds,* Oxford University Press, New York.

Gill E. K. (1993), Long Distance Presentations, *Presentation Products Magazine,* May 1993, pp. 19-26.

Harrison T.M. and Stephen T. (1992), On-Line Disciplines: Computer-Mediated Scholarship in the Humanities and Social Sciences, *Computers and the Humanities,* 26, pp. 13-25.

Kaye A.R. (1992), Computer Conferencing and Mass Distance Education, *Empowering Networks: Computer Conferencing in Education*, Educational Technology Publications Inc., Englewood Cliffs, New Jersey, pp. 69-100.

Lewis C.T. and Hedegaard T. (1993), Online Education: Issues and Some Answers, *Technological Horizons in Education*, April 1993, pp. 68-71.

McCormick N.B. and McCormick J.W. (1992), Computer Friends and Foes: Content of Undergraduates' Electronic Mail, *Computers in Human Behavior*, 8, pp. 379-405.

Randall J. P. (1992), The Emerging Potential of Virtual Reality in Postsecondary Education, *New Directions for Teaching and Learning*, 51, pp. 77-81.

Reingold H. (1991), *Virtual Reality*, Simon and Schuster, New York.

Riel M. (1990), Cooperative Learning Across Classrooms in Electronic Learning Circles, *Instructional Science*, 19, pp. 445-466.

Riel M.M. and Levin J.A. (1990), Building Electronic Communities: Success and Failure in Computer Networking, *Instructional Science*, 19, pp. 145-169.

Ronkin R.R. (1991), *Global Cyberspace: Who Needs It?* Network communication: Use anonymous FTP to dhvx20.csudh.edu, filename: global-net/global-cyberspace.txt.

Roupp R. (1993), Agenda for an Uncertain Future, *LabNet: Toward a Community of Practice*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 291-310.

Roupp R. and Pfister M. (1993), An Introduction to LabNet, *LabNet: Toward a Community of Practice*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 1-20.

Shapard J. (1990), Observations on Cross-Cultural Electronic Networking, *Whole Earth Review*, Winter 1990, pp. 32-35.

Snee R. (1993), What's Missing in Statistical Education?, *The American Statistician*, 47, pp. 149-154.

Utsumi T. and Villarroel A. (1992), *Towards Establishing a Global/Latin American (Electronic) University*, Network communication: Use anonymous FTP to dhvx20.csudh.edu, filename: global-net/global-univ.txt.

Van Deerlin L. (1980), Applying Commercial Communications to Social Services, *World Progress Through International Telecommunications*, United Nations Association of the United States of America, Inc.

Wegner L.S. (1992), The Research Library and Emerging Information Technology, *New Directions for Teaching and Learning*, 51, pp. 83-90.

COGNITIVE TECHNOLOGIES FOR STATISTICS EDUCATION: RELATING THE PERSPECTIVE OF TOOLS FOR LEARNING AND OF TOOLS FOR DOING STATISTICS

Rolf Biehler
Institut für Didaktik der Mathematik, Universitat Bielefeld
Postfach 100131, D - 33501 Bielefeld, Germany

1. The software dilemma in statistics education

For introductory statistics education, several types of software are relevant and in use: custom designed educational programs for a specific educational goal, statistical systems for data analysis (in full professional version, in student version or as a specifically designed tool for students), statistical programming environments, spreadsheets and general purpose programming languages. On both a practical and a theoretical level, we can perceive a double dilemma, which is the worse the lower the educational level we have in mind. On the one hand, we have professional statistical systems that are very complex and call for high cognitive entry costs, although they flexibly assist experts. On the other hand, custom designed educational software is of necessity *constrained* to enable students to concentrate on essential aspects of a learning situation and to make likely certain intended cognitive processes. Nevertheless, as these *microworlds*, as we will call them here for short, are often not adaptable to teachers' needs, they are often criticized as being too constrained. Their support for flexible data analysis is limited, and to satisfy the variety of demands one would need a collection of them. However, coping with uncoordinated interfaces, notations and ideas in one course would overtax the average teacher and student. This practical dilemma is reflected on a theoretical level. It is not yet clear enough what kind of software is required and helpful for statistics education. We need a critical evaluation and analysis of the design and use of existing educational and professional programs. The identification of key elements of software that are likely to survive the next quantum leap of technological development and that are fundamental for introductory statistics is an important research topic. Results should guide new "home grown" developments of educational programs or, facing the difficulty of such developments, should influence the adaptation and elaboration of existing statistical systems toward systems that are also more adequate for purposes of introducing and learning statistics.

We might imagine ideal software support for introductory statistics education as follows:

- a student tool for data analysis and for modelling that can grow and expand into a professional version on several paths – instead of mere technically reduced student versions;
- a system of co-ordinated computer experiments, learning environments and major visualisations – that can be adapted to students' and teachers' needs;
- a larger host system that contains the student tool and the learning environments as integral parts. By that it also guarantees some notational and conceptual coherence and cognitive economy between learning environments and tools for doing statistics. The host system can be used as a tool-maker or medium builder for statistics education, at least in the hands of average experts.

These goals and research issues provide a framework that is to be filled with content. We will give some ideas and directions that are partly based on the results of two projects. In the "Software tools for statistics"–project, we explored the potential of commercially available software tools for statistics from an educational point of view (see Biehler and Rach, 1990a, Biehler and Rach, 1990b). In the ensuing MEDASS project, we have used the results to develop a detailed paper-and-pencil prototype of a software tool for teaching and learning data analysis. The detailed specification is described in Biehler and Rach (1992), the didactical background in Biehler (1992).

2. Cognitive technologies and the meaning of the mean

The notion of cognitive technology has induced us to discuss computer software from a broader perspective. Written language, symbol systems (mathematical and others), films, pictorial media, and computer languages can be interpreted as cognitive technologies. Pea (1987, p. 91) describes cognitive technologies as any "medium that helps transcend the limitation of the mind ... in thinking, learning, and problem solving activities", and he underlines that cognitive technologies "have had remarkable consequences on the varieties of intelligence, the functions of human thinking, and past intellectual achievements". Computer software can be regarded as a new kind of cognitive technology that also has assimilated and transformed some traditional technologies like graphs and formal notations. We may regard education as a system that does not just transmit knowledge but also "tools of the intellect" provided by the culture to enhance students' thinking. The question of what software is adequate for statistics education can be reformulated to ask which cognitive tools for statistics should be provided.

How can cognitive technologies contribute to understanding statistical concepts? In the following, we shall take the concept of mean as an example. The meaning of concepts can be characterised by the relations in an epistemological triangle in Fig. 1, which we adapt for our purposes from Steinbring (1991). For example, the concept of mean acquires its meaning from its relation to other concepts such as distribution, median etc., from representations and tools to work with it, such as formulas, algorithms, visualisations and from its applications to solving statistical problems, from its interpretations in contexts of application.

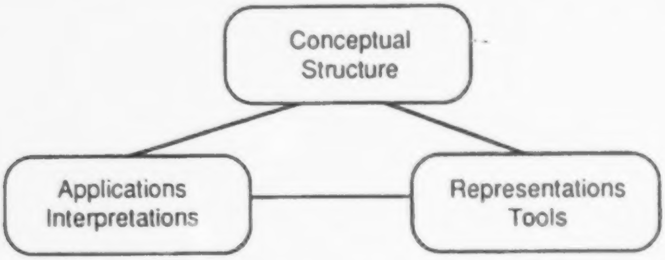


Figure 1. *Epistemological triangle*

Some empirical research, e.g. Pollatsek *et al.* (1981), confirms the impression that students often have a mere computational (procedural) understanding of the mean. In other words, the relational structure of Fig. 1 is an ideal seldom achieved. Important aspects of the mean are a geometric interpretation, i.e. its interpretation as a "center of gravity", and its varying significance as a summary statistics with regard to different types of data distributions. This is particularly important in comparison with other summary statistics for the center of a distribution. Furthermore, a functional interpretation is relevant, for instance, for understanding non-resistance to outliers: the mean is a function of the data and will change in a certain way when the data change.

New cognitive technologies obviously offer new representations providing didactical opportunities. In the case of summary statistics, we find the suggestion to use spreadsheets: When the data have been put into a range of cells, some further cells can be defined as the mean, the median or other functions of the data. Changing the data "by hand" is easy in a spreadsheet and will automatically lead to an update of the summary values: the "function machine" behind a spreadsheet is a new representation of the functional dependence. In particular, it supports an experimental study of the functional relationship by students. Whereas this example uses a general purpose tool, we also find more constrained and closed learning environments with new representations. The option *Stretchy Histogram* from the software *Statistics Workshop* provides opportunities for an interactive "geometrical experiment" in statistics education (see Fig. 2). The user is able to manipulate the shape of the histogram by directly pulling a bar in the graph. The mean, the median and the two quartiles are functionally dependent on the distribution with automatic updating. The relationship between the shape of a distribution and the four summaries can be explored empirically. The qualitative new feature of *direct manipulation* is used in this program. We could also transpose numerico-algebraic experiments from a spreadsheet environment into geometrical experiments. Imagine that you can move data points in a dot plot and observe how a graphical display of summary statistics, say a boxplot, will change.

One could also think of transforming other relevant visualisations of the summary statistics, as e.g. are discussed by Bentz and Borovcnik (1984), into interactive, dynamical geometrical computer experiments. For an empirical analysis of the pedagogical payoff of such environments it would also be helpful if a variety of visualisations were available. However, *Stretchy Histograms* is closed and too constrained and would not support these variations.

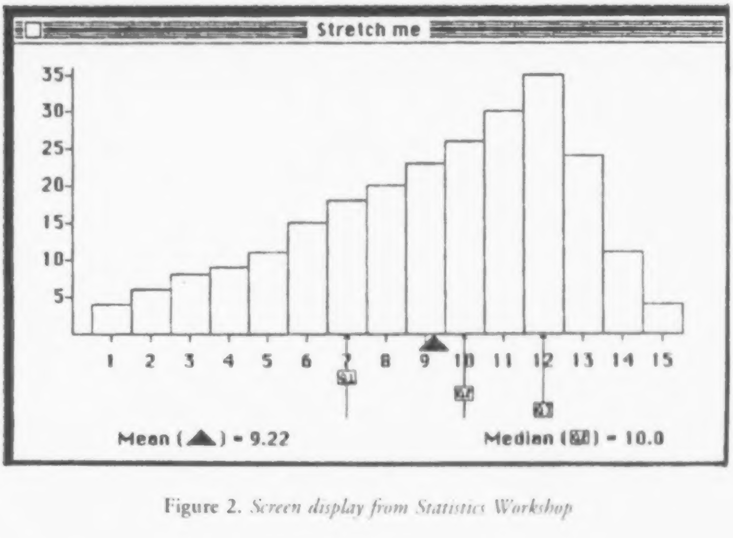


Figure 2. *Screen display from Statistics Workshop*

Due to new technological developments, the status of certain cognitive technologies may change. Algebra has always been a powerful cognitive technology for specifying how things have to be calculated and for depicting relations: when computers do the calculating, the latter becomes even more important. For instance, "experts" can "see", in the formal definition of the mean, that the mean is a linear function of each of the data values. Being able to interpret algebraic notation is still very relevant. Visual and experimental opportunities should be used to enhance this algebraic proficiency instead of aiming at its substitution.

Cognitive technologies affect the other elements of the epistemological triangle as well. From the perspective of teaching applied statistics, applications and interpretations of the mean are important constituents for understanding. Statistical software tools together with resources of statistical data make accessible qualitatively new kinds of applications and experiences while using the mean and other summary statistics in data analyses. For instance, students may re-discover the value of summarizing data to recognize lawlike relationships, whereas chaos and randomness prevail without summarization. This was a basic experience of the 19*th* century social statisticians when more data became available to them. The new technologies could turn this into an experience for students. Technology may support the sharing of rich data analytic experience that has been a basis for conceptual and methodological developments in history. For instance, different interpretations of the mean from substantialist interpretations as a "true value" to its more relational interpretations as a method for comparison or even as no more than a formal model parameter could be experienced in different applications.

The tension we have established between using statistical software tools and using educational programs is less extreme than it looks. One reason for developing *Stretchy Histograms* is certainly that currently available statistical systems will not support the above experiment. However, systems with properties that allow such more constrained environments to be defined could be also of use for the working statistician. For instance, being able to enrich the display of a data distribution by marking summary statistics would be also useful in data analysis, although many current systems do not support this as a general feature. Furthermore, the possibility of directly manipulating distribution shape could be a useful feature for a research statistician in times where the relative importance of algebraically defined curves has decreased in favour of *non-parametric* densities and functional relations. We shall come back to the possibility of using statistical systems as host systems for custom designed educational programs later.

3. Cognitive technologies as reorganizers

Cognitive technologies are frequently regarded as "cultural amplifiers" of the intellect. That means a quantitative intensification without a qualitative change of structure. However, Pea (1987) emphasizes the potential of cognitive technologies for leading towards a qualitative reorganization of thinking. There is great promise in the analysis of the educational possibilities of cognitive technologies from the reorganizational point of view. What may be called reorganization of thinking on an individual cognitive level may be called a scientific revolution on the level of science. Obviously, statistics provides good examples for this thesis. Let us take the development of Exploratory Data Analysis (EDA) as an outstanding example. Computer technology does not only support doing more analyses and graphs. The possibility of multiple analyses and interactive graphics has also given rise to qualitatively new types of data analyses and of theoretical research in statistics.

Examples of qualitatively new aspects of statistical thinking include:

- employing graphical displays as analytical tools;
- using concepts and methods for data analysis without a probabilistic justification;
- overcoming oversimplified model assumptions: diagnostic checking, data guided model choice, robust methods, effective handling of models with weaker (broader) assumptions;
- iterative, interactive ways of analysing data;
- exploration and insight as goals of data analysis, in addition to inference;
- resampling techniques as an alternative to classical inference methods.

Just as we can observe that certain statistical software systems have been an important (necessary but not sufficient) factor for making these ways of thinking available to a broader group of people, we have to think of tools for introducing statistics also in terms of their potential for supporting these new qualitative dimensions.

There is another important aspect in the development of statistics. The above changes of the objects of research have also affected the methods of research in statistics. It is said that statistics has become more like an experimental science: statisticians construct new methods and analyse them not only by mathematical (analytical) methods but also by sophisticated simulation studies and by trials on sets with real data. These research methods may also be more accessible to students than highly theory-laden analytical methods. Students might participate in a qualitatively new way of thinking (from their perspective), namely at the statistical workbench of constructing and evaluating new methods, instead of being only confronted with ready-made methods. For instance, students may suggest their own test statistics, which can be studied easily by simulation under different hypotheses. Or, students may design new graphs and test them on real data with real people.

How can we understand the cognitive role of concrete tools in more detail? There are three relevant features of software tools: "as user-directed agents that perform actions for the user"; as "built-in constraints or supports", as Kaput (1992, p. 525) expresses it, and as new kinds of "representations and objects to think with". The latter two are properties of a system as a whole, and not of individual agents and procedures.

A problem is the large variety of tools whose design has interacted with statistical practice. There is no agreement on standards for software tools with regard to certain domains of application and styles of thinking. Depending on their constraint-support structure, existing statistical systems are not equal in supporting the above mentioned qualitative ways of thinking. It is very likely that differences will show up between regular *SPSS*, *S-Plus* and *Data Desk* users. Furthermore, the same concepts can have specific meanings within the context of every particular software. For instance, a "table" is different in software such as *StatView*, *MINITAB* or in an ordinary spreadsheet, because there are different internal representations and different possibilities of manipulating a table; these also make electronic data tables different from tables on a sheet of paper. Another example are statistical graphs which are manipulable objects dynamically linked to the data in *Data Desk*, which can be defined and edited by scripts in *S-Plus*, and are mere bitmap or vector graphics in many educational programs. Although this shift of meaning in relation to the new objects and operations in a software tool, designated by Balacheff (1993) as the *computational transposition*, can be a source of misunderstanding, their potential for new types of activities has to be underlined, Dörfler (1993).

In the following paragraphs, we will critically examine some typical features of tools for learning and doing statistics in educational contexts.

4. Tools for data analysis as host systems for learning environments

The idea to co-use systems for statistical data analysis for building educational microworlds is not new. *MINITAB* has often been used in this sense by selecting a limited subset of commands for simulation and experimentation. The more coherent way of extending the language and the possibility of constructing new graphs makes *S-Plus* a better candidate for defining statistical microworlds, however at the price of much higher cognitive entry cost. A didactical proposal to use the programming language *APL*, extended by appropriate statistical functions for such purposes, is described in Naeve *et al.* (1991). Spreadsheets have also been considered as host systems. A recent reference is Arganbright (1992). As tools for flexible data analysis current spreadsheets have important weaknesses. Because their concepts and in-built functions are not well adapted to statistical concepts, data organisation and management may become awkward.

For secondary education, but also for introductory statistics in higher education, starting with a command language system is problematic. We are convinced that a host system with a graphical user interface and dynamically linked windows offers a more adequate basis. This would be even more the case if the limits of menu-based systems could be mitigated by allowing *formal language input* at those places where there is a real need, for instance, for defining data transformations or new procedures. An essential point is, how easy it is to define new micro-worlds, whether programming expertise is needed or whether we can do "programming by example". Two trends in software maturation meet these requirements: the general effort of providing easier ways of customization and the trend of statistical tools to support more open explorations.

The software *Data Desk* offers quite interesting possibilities for defining microworlds and we will discuss possibilities and limitations. *Data Desk* supports the definition of nested "derived variables" by formula, the linking of representations (data, graphs, equations of derived variables) and the possibility of saving environments with empty data or pseudo-variables (conserving all the defined relations).

In the screendump in Fig. 3, we have a set of car data where we have made a regression of *MPG* (miles per gallon) vs. *weight* of the car (window 4). As the graph has curvature, we would like to try experimentally a transformation of *MPG* from the family of transformations x^m. *Data Desk* offers *sliders* for an easy direct manipulation of parameters. We define a "slider" with name *m* (window 1), that is a parameter that can be directly maniputated by moving the vertical line on the screen. We define a "derived variable" *f(MPG)* = MPG^m (window 2). Window 3 shows a regression line of *f(MPG)* vs. *weight*, window 5 and 6 show two displays of residuals of that regression. If we change *m* in window 1, all the dependent windows will update. The dependence structure can be summarized as in the right part of Fig. 3.

This microworld can be used for studying the effect of different data transformations. One could also use artificial data or data from a certain known model to study these relations. In statistical practice, this microworld could be used for an interactive choice of a linearizing transformation – instead of using ready-made automatic methods for this purpose. As it is possible to save this environment and to replace the variables chosen by different ones, we have really defined a small new microworld for exploring the effects of data transformation. It would have also been possible to directly edit the formula in window 2. We note that formal-mathematical representations have re-entered at a reasonable location in a limited way.

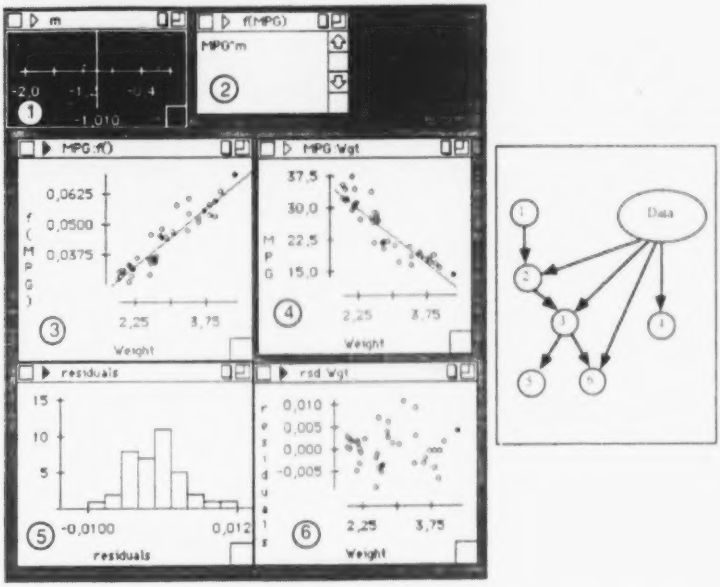

m
-2,0
-1,4
-0,4
-1,010
①
f(MPG)
MPG^m
②
MPG:f()
0,0625
0,0500
0,0375
f(MPG)
2,25
3,75
Weight
③
MPG:Wgt
37,5
30,0
22,5
15,0
MPG
2,25
3,75
Weight
④
residuals
15
10
5
-0,0100
0,012
residuals
⑤
rsd:Wgt
0,010
0,005
0,000
-0,005
residuals
2,25
3,75
Weight
⑥
Data
1
2
3
4
5
6

Figure 3. *An experimental environment in Data Desk 3.0; the linkage structure*

This example shows how new requirements from practice meet requirements from learning environments. In *Data Desk*, however, it would be impossible to define the standard visualisation of confidence intervals, namely simulating a sequence of samples and displaying the (random) sequence of pertinent confidence intervalls in one display.

We would need to define an experiment in the following steps:

(1) choose a probability *model* with expected value μ;
(2) draw m samples of size n;
(3) calculate the m means;
(4) chose the level $(1-\alpha)$;

(5) calculate the interval lengths according to a method *K*;
(6) plot the *m* means vs. 1, ... , *m*;
(7) add symmetric bars to the scatterplot, corresponding to the confidence intervals;
(8) add a dotted line at $y = \mu$.

These steps should be executed by using commands from the menu system. They should result in new objects (data or plots) that are dependent of the previous objects and that can be represented in separate windows. In Data Desk, steps 1-6 would be possible for some models; however, the resulting objects would not be dependent in the required sense. The necessary enrichments of the display (7) and (8) are not possible. Now, the user might like to control and change the parameters *model*, μ, m, n, $(1-\alpha)$, K. Also, it should be possible to use another random seed, that is repeating the whole random experiment in a simple way, for instance by pressing a "recalculation button". The final plot with the series of confidence intervals must dynamically *depend* on these parameters. This is not the case in Data Desk. An adequate host system should provide better dependence structures. We have made some suggestions for that in Biehler and Rach (1992). The repeated sampling principle which should be visualized by the above experiment could be similarly demonstrated by applying the confidence method to a series of real data sets, as in Biehler (1990) to the famous historical data on boys and girls christened in London from 1629 to 1710. This may contribute to meaning in the sense of applications and interpretations in Fig. 1. The similarity of both visualisations supports also the need for integrating data analysis tools and learning environments.

5. Student tools for simulation and modelling

The software dilemma is especially grave with regard to elementary tools for simulation and modelling, as was described in more detail in Biehler (1991). Apart from statistical programming environments, statistical systems for data analysis are weak in this respect. *Models* are not separate objects of the systems. It is difficult to generate the requested random data, especially with simple multistage experiments, and data analysis is not well adapted to thinking in terms of *events* and *random variables* whose frequency or distribution has to be recorded during simulation. Educational programs with ready-made simulations are valuable but will probably not contribute to involving students in the process of modelling. If we are interested in this, it will be essential that the students themselves define the model as well as transform the verbally described events and random variables into a more abstract notation.

The educational software *Probability Simulator* is an important step towards closing this gap by means of a menu-driven multi-window environment with many relevant applications in elementary probability education (see Konold, 1991). The students can:

(1) choose and change a box model;
(2) fix sample size and number of repetitions;
(3) run and re-run the simulation;
(4) define (several) events that are to be counted.

The results of counting the events are displayed in a separate window which is dynamically linked to the data window containing the raw data. The language for defining events is a delicate problem. *ProbSim* uses wildcards (*) and *variables* (v_1, ...) for that purpose and makes a distinction between *ordered* and *unordered* events. For instance, the event "at least one pair of equal outcomes in an experiment of sample size 6" is defined by the *unordered* type of event (v_1,v_1,*,*,*,*), the *ordered* type would mean that at least one pair occurs in the first two places. As compared to formal programming languages, students can define models in terms of urns and define events in a simpler albeit abstract notational system. The constraint-support structure of *ProbSim* corresponds to a certain idea of what should be essential steps in model building for secondary probability education. Every student tool should have at least similar properties.

However, it is not (yet) possible to define random variables and record their distribution. Also, there is an abundance of simple multi-stage experiments in elementary probability with dependent trials that cannot be simulated because *ProbSim*, as nearly all similar programs, is confined to independent repetitions. After some experience, one would also appreciate the availability of a more general language for defining events by the user. These are some of the possibilities for expanding a student tool towards wider educational applications and professional requirements. Disparate development projects and constrained resources for educational software development have not yet led to overall satisfying solutions.

6. Student tools for data analysis

An obvious criterion for tool design is making "simple things simple". The interpretation of this depends on those intended applications of the tool that should be simple to work on. There are some recent student tools for data analysis that have such a profile dependent on a certain curriculum or a didactical intention, namely *DataScope*, *Statistics Workshop* and *TableTop*.

In Biehler and Steinbring (1991), we have described experience with the GRAPHDAS pilot project in introducing ideas of data analysis in the classroom. We used *StatView* and *Data Desk* in our exploratory work with teachers and students, and found the moderate openness of these tools helpful in two respects. Firstly, we needed this openness as we did not want to make strong assumptions beforehand concerning what kind and content of data analysis would be adequate for the secondary level. Using teaching experiments for exploring this problem could not have worked with less open, curriculum dependent, tools. Secondly, we assumed that some small project activity with open problems referring to complex data should be part of a course. In other words, we included problems that could lead to new methodology, at least from the students' perspective – sometimes from that of teachers and from our own. For instance, although teachers introduced data on traffic accidents as a topic where boxplots should be used to analyse seasonal variation, students asked for predictions of future developments and the teacher found himself in a situation of time series forecasting. If data analysis is to be stimulated and controlled by students' subject matter interests and knowledge, a tool's closedness must not crush such involvement. From this perspective, the above curriculum dependent tools were not adequate for our purpose. The professional tools, however, would have been much more useful, especially as tools for student work, if they had provided greater adaptability. For instance, an easy redesignability of the menu system would have been very valuable. Nevertheless, there was not one tool that fulfilled all the elementary functions we found useful in the project work. This claim can be illustrated with regard to statistical graphs. No tool offered the direct interaction with graphs of *Data Desk*, the composite graphs of *StatView*, the variety of elementary graphs of *Systat*, and the possibility of enhancing graphs by further statistical information of *S-Plus* – not to speak of *S-Plus*' capability to define new graphs.

A design or evaluation of software cannot be performed neutrally, but only with regard to a class of intended applications for which the software should provide easy solutions. An explicit discussion of systems of intended key applications may contribute to a more objective judgment. We shall discuss some intended applications in the following, where we found that existing software was not very strong, but which are relevant from our view of elementary data analysis. A more comprehensive account can be found in Biehler and Rach (1992) and Biehler (1992).

Categorical and nominal variables. Let us take as an example a data set with *height* and *weight* measurements for a group of male and female students coming from five age groups. We represent the *weight* for each student in a bar graph labeled by the students' names. Usually, an alphabetical order of the names or the underlying index numbers is taken as the basis for the plot. Re-sorting the names in the plot is an elementary method of analysis. Now, we would like an easy access to the following options: switch from barchart to dotchart in the sense of Cleveland (1985); project the points of the dotchart to an axis to explicitly show the marginal distribution as a dotplot; re-sorting the names according to the variable *weight*, re-sorting according to the variable *height* as a means to approach association of these two variables, regrouping the names according to a categorical variable like sex and then sorting within the categories to produce a grouped dotchart or bar chart. We can conceive of these operations as manipulating a categorical or nominal axis and not as transforming the data into auxiliary data first and then replotting them by re-calling a graphical procedure. We might generalize this idea to plotting two categorical variables against each other. This results in a "scatterplot with categorical or nominal axis". The latter idea is realized in the software *TableTop* (see Hancock and Kaput, 1990). If we now had the opportunity of re-sorting and permuting the order of the 2 categories on each axis "by hand", we might discover new relationships and possible classifications. In fact this would be a way of implementing the graphical method of Bertin (1977).

Raw and aggregated data. The computer has greatly extended the possibilities of data analysis with raw data. However, there are still many important societal (very large) data sets which are only available in aggregated form. It should be easy to work with the same statistical methods on either data type (with slightly different interpretations of the commands). For instance, although it is often possible to vary the length of a class interval, when calling a histogram to display a distribution for ungrouped data, this possibility should be also available for grouped data. In the GRAPHDAS project where we worked with children and teachers of grade 8, we had data on the number of traffic accidents per age group of length 1. For a comparison of the fairly irregular curves of males and females, a flexible regrouping to smooth the two histograms would be helpful.

Analysis by groups should be a simple possibility. For some statistical methods, this is realized in educational programs like *Statistics Workshop* and *DataScope*. *DataScope* makes it also easy to select two grouping variables to base analyses on a two-way classification. *Data Desk* shows a fairly general realisation, where we can have graphical displays and numerical summaries by group, although no easy two-way classification. Also, for a small number of category levels one would like to have not only many windows for each group but integrated graphs. For instance, selecting a stem-and-leaf display by group (of two categories) should result in a back-to-back stem-and-leaf.

Composition of graphs. The juxtaposition of graphs should be an easy possibility. One might wish for some "intelligence" in juxtaposing graphs, for instance with regard to identifying or adapting axes of the same variable. The superposition of graphs or the composition and enrichment of statistical graphs should become as easy as in modern drawing programs, but with the essential difference that the graphical elements can be mathematically or statistically defined and processed.

Graphs as an interface to data: geometrization of data analysis. Velleman (1989) employs the metaphor of using graphs to talk to the data. We may think of options for a geometrical classification into groups "by hand", selecting subsets and displaying information of a data case represented by a point in a graph (identification). This has been realised in some pieces of recent software. However, having elementary educational applications in mind, we may also think of possibilities of graphical input of data (to generate a data set with specific geometric properties, realized in *Stat-Lab*), and of graphical change of data (graphical dragging of points to study how analyses depend on a data point, similar to the dragging mode in the geometry software *Cabri Géomètre*). Movable lines that can be drawn by hand into a data set, should not only be didactical means to prepare the method of least squares as in *Statistics Workshop*, but can be an implementation of the method of "eye-fitting" lines to data. However, it should be as easy as with least squares lines to get residuals from such lines. It should be equally easy to draw and manipulate a stepwise linear curve "by hand" into a set of points, and then work with this line as with other statistical summaries. A slicing of a scatterplot into a number of stripes (graphical discretization), then calculating summary points to get a rough smooth is an elementary smoothing method that has the advantage of being transparent to the user with regard to its construction – as compared to more advanced methods.

7. Conclusion: Requirements from an ideal softwaretool for supporting learning and doing statistics

What features of a software for supporting learning and doing statistics are important? Let us summarize the discussion in some key points:

(1) The software should possess a *graphical user interface* with multiple, dynamically linked interactive representations (graphics, tables, formulas, analysis results).

(2) The limits of menu-based systems should be mitigated by allowing *formal language input* at those places where there is a real need, for instance for defining data transformations or "derived variables".

(3) The software should have the character of a *meta-tool* or *meta-medium*, which allows the definition of more constrained exploratory learning environments or more constrained tools that are better suited for certain users and certain domains of application.

(4) The software should not only provide ready-made methods, but support the *definition or customisation of new and personal, non-standard methods*, for instance graphs, tables, statistical tests. This is optimally realized by "*programming by example*" to avoid programming in formal languages.

(5) The software should support an easy access and a flexible use of *elementary* graphical and numerical methods of *exploratory and inferential data analysis. Bootstrapping* should be "at hand" for analysing the variability of summaries.

(6) The software should support the *exploration of statistical methods* on real, simulated and artificially constructed data sets. Exploring the changes in results when changing data or method parameters should be easy.

(7) The software should support the *definition of computer experiments* to explore the consequences of parameter changes as well as of chance variations.

(8) The *construction of visualisations* should be supported by something like a graph construction set.

(9) The *definition of statistical and probabilistic models* should be supported. It should be possible to define events and random variables of interest, and to analyze the simulated data like real data.

Of necessity, these points define a framework that has to be elaborated further. In any case, we should be motivated as statistics educators to develop our own perspective and needs, hoping that we can get intellectual and material support to improve the state of the art.

SOFTWARE

Cabri Géomètre. Laborde J. M., *et al.*, LSDD, B.P. 53 X, F-38041 Grenoble-Cedex (Macintosh, MS-DOS).

Data Desk 3.0 (1989), Velleman P., Data Description Inc., P.O. Box 4555, Ithaca, NY 14852 (Macintosh).

DataScope (1992) version 1.16. Konold C. and Miller C.D., Scientific Reasoning Research Institute, Hasbrouck Lab, University of Massachussetts, Amherst, MA 01003 (Macintosh).

MINITAB. MINITAB Inc., 3081 Enterprise Drive, State College, PA, 16801 USA (MS-DOS, Macintosh and others).

ProbSim - A Probability Simulator. (1992) version 1.3.2, Konold C. and Miller C.D., Scientific Reasoning Research Institute, Hasbrouck Lab, University of Massachussetts, Amherst, MA 01003 (Macintosh).

S. and *S-Plus* (1988), Becker R.A., Chambers J.M. and Wilks, A. R., see: Becker, Richard A., Chambers, John M., Wilks, Allan R., The New S Language, Wadsworth & Brooks, Pacific Grove. (Unix, MS-DOS/386 and others).

SPSS. SPSS Inc., 444 North Michigan Avenue, Chicago, IL 60611 (Macintosh, MS-DOS and various other systems).

Statistics Workshop (1991), Rubin A. and Bruce B. (BBN Inc., Cambridge, MA, USA). Wings for Learning/Sunburst, 1600 Green Hills Road P.O. Box 660002, Scotts Valley, CA 95067-0002 (Macintosh).

StatLab (1987). W. Douglas Stirling. The New Zealand Statistical Association, P.O. Box 1731, Wellington, New Zealand (Macintosh)

StatView SE+Graphics (1988), Feldman D. S. jr and Gagnon J., Abacus Concepts, Inc., 1984 Bonita Ave., Berkeley, CA 94704, USA. (Macintosh).

StatView 4.0 (1992), Feldman D.S. and Gagnon J., Abacus Concepts, Inc., 1984 Bonita Ave., Berkeley, CA 94704, USA (Macintosh).

Systat 5.0 (1989), SYSTAT Inc., 1800 Sherman Avenue, Evanston, IL 60201-3793 (Macintosh, MS-DOS).

TableTop, software in development, Hancock C. and Kaput J.J., TERC, 2067 Massachusetts Avenue, Cambridge MA 02140.

Bibliography

Arganbright D. E. (1992), Using spreadsheets in teaching statistics, in F. Gordon and S. Gordon (eds.), *Statistics for the Twenty-First Century. MAA Notes n. 26*, pp. 226-242, Washington, DC, Mathematical Association of America.

Balacheff N. (1993), Artificial intelligence and real teaching, in C. Keitel and K. Ruthven (eds.), *Learning from Computers: Mathematics Education and Technology*, Springer, Berlin, pp. 159-186.

Bentz H.J. and Borovcnik M. (1984), Mittelwert, Median und Streuung: Eine Zusammenschau, in H. Kautschitsch and W. Metzler (eds.), *Anschauung als Anregung zum mathematischen Tun: 3. Workshop zur 'Visualisierung in der Mathematik' vom 11. - 16. Juli 1983 in Klagenfurt*, pp. 208-220, Hölder-Pichler-Tempsky, Wien.

Bertin J. (1977), *La Graphique et le Traitement Graphique de l'Information. [Engl. Transl.: Graphics and Graphic Information-Processing, Berlin: de Gruyter 1981]*, Flammarion, Paris.

Biehler R. (1990), Daten analysieren mit dem Computer: Unterstützung von Begriffsbildung und Anwendungsorientierung in der Stochastik, *Der Mathematikunterricht*, 36(6), pp. 50-71.

Biehler R. (1991), Computers in probability education, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters - Probability in Education*, pp. 169-211, Kluwer, Dordrecht.

Biehler R. (1992), *Intendierte Anwendungen und didaktische Begründungen zu einem Softwarewerkzeug zur Explorativen Datenanalyse und stochastischen Simulation für Schule und Ausbildung*, Inst. f. Film und Bild in Wissenschaft und Unterricht (FWU), München.

Biehler R. and Rach W. (1990a), Softwaretools for statistical data analysis in education and teacher training? An analysis of software conceptions from an educational perspective, in F. Faulbaum *et al.* (eds.), *SOFTSTAT '89: Fortschritte der Statistiksoftware 2*, pp. 575-583, G. Fischer, Stuttgart.

Biehler R. and Rach W. (1990b), *Softwaretools zur Statistik und Datenanalyse: Beispiele, Anwendungen und Konzepte aus didaktischer Sicht*, Soester Verlagskontor, Soest.

Biehler R. and Rach W. (1992), *MEDASS: Explorative Datenanalyse und Stochastische Simulation - Anforderungsbeschreibung zu einem Softwarewerkzeug für Schule und Ausbildung*, Institut für Didaktik der Mathematik, Universität Bielefeld, Bielefeld,.

Biehler R. and Steinbring, H. (1991), Entdeckende Statistik, Stengel-und-Blätter, Boxplots: Konzepte, Begründungen und Erfahrungen eines Unterrichtsversuches, *Der Mathematikunterricht*, 37(6), pp. 5-32.

Cleveland W.S. (1985), *The Elements of Graphing Data*, Wadsworth, Monterey.

Dörfler W. (1993), Computer use and views of the mind, in C. Keitel and K. Ruthven (eds.), *Learning from Computers: Mathematics Education and Technology*, Springer, Berlin (to appear).

Hancock C. and Kaput J.J. (1990), Computerized tools and the process of data modeling, in G. Booker, P. Cobb and T.N. d. Mendicutti (eds.), *Proceedings of the 14th International Conference on the Psychology of Mathematics Education*, 3, pp. 165-172, The Program Committee of the 14*th* PME Conference, Mexico City.

Kaput J.J. (1992), Technology and mathematics education, in D. Grouws (ed.), *Handbook of Research on Mathematics Teaching and Learning*, pp. 515-555, Macmillan, New York.

Konold C. (1991), ChancePlus: A Computer Based Currriculum for Probability and Statistics. *Second Year Report*, Scientific Reasoning Research Institute, University of Massachusetts, Amherst.

Naeve P., Trenkler D. and Wolf H.P. (1991), How to make the teaching of statistics roar - Some thoughts on computer based experiments. *Computational Statistics Quarterly*, 6(4), pp. 325-353.

Pea R.D. (1987), Cognitive technologies for mathematics education, in A.H. Schoenfeld (ed.), *Cognitive Science and Mathematics Education*, pp. 89-122, L. Erlbaum, Hillsdale.

Pollatsek A., Lima S. and Well A.D. (1981), Concept or computation: students' understanding of the mean, *Educational Studies in Mathematics*, 12, pp. 191-204.

Steinbring H. (1991), The theoretical nature of probability in the classroom, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters - Probability in Education*, pp. 135-167, Kluwer, Dordrecht.

Velleman P.F. (1989), *Handbook of Data Desk 3.0*, IL., Odesta Corporation, Northbrook.

USING COMPUTERS IN TEACHING STATISTICS IS IT WORTHWHILE?

Andrej Blejec
Institute of Biology, University of Ljubljana
Karlovska 19 POB 141, 61001 Ljubljana, Slovenia

1. Introduction

Computers were invented and constructed to compute. Since statistical analyses are computing intensive it is natural that computers are widely used in statistical research and applications. Statistical applications, such as census with its problems of sorting, counting and tabulating, were among the motives for constructing the ancestors of modern computers.

It was always clear that students have to be taught how to use computers, because they will use them in their later careers. With the development of modern technology, computers evolved from large mainframes to personal computers, available for individual use. Availability of personal computers changed the way of computer usage and allowed computers to be incorporated into the teaching process in various disciplines.

In particular, can computers be useful in teaching of statistics? If yes, to what extent? What changes in the teaching process are needed if we want to apply computers efficiently? There are many other questions related to the usage of computers in teaching. I would like to present some of my views about computers in teaching of statistics.

Computers can be used in teaching in many ways. Some are of a practical nature, speeding up routine actions such as calculating or sorting. Some help students to understand, and teachers to explain, certain difficult topics. Computers can be used as electronic textbooks. They can be incorporated into practical student work as well as into lectures.

2. Computation

One aspect of the use of computers in teaching of statistics is the computer as an aid to calculation. The benefit of different types of calculators in the past, and computers at the present time, is enormous. Computers take the burden of computing from users, leaving them time for final decisions based on statistical calculations. I wrote users, not students, because the burden of computing certainly falls on users of statistics. But what about students? Is it wise to take *all* calculations from students in the learning phase? I don't think it is. Some personal experience and involvement in the calculation process can help to better understand the formulas and statistical methods considered. Such calculations *by hand* should certainly be limited to small datasets, which are quite common in teaching.

Moderate or large datasets, more realistic for statistical decision-making, cannot be analysed without computers. In the past, the calculation efforts were too large and reasonably large data sets were seldom analysed by students. What software should now be used: statistical packages, spreadsheets or both? To be on the safe side it is wise to answer-both. Let us first consider spreadsheets.

There are some advantages in using *spreadsheets* in teaching statistics. Using computers, people are sometimes confused by the fact that data are hidden in a mystery called "file", where they are kept invisible. In spreadsheets, data are kept in tabular form, similar to the hand-written tables used for data recording data on paper. Formulas, necessary for calculations, are entered step by step into individual cells, showing the intermediate results of calculations. Many statistical functions are implemented in modern spreadsheets, from average values to inverse probability functions, regression analysis and hypothesis testing. For teaching purposes, the sum of a range of cells is the crucial one, and always present. It can be clearly shown that average value, variance and many others are indeed functions, defined on a sample (or realisation of a random variable, process...). Entering statistical formulas in fragments, showing the steps of the calculation scheme on actual data, can aid understanding of statistical formulas. The automatic recalculation facility can be used to experiment with "what-if" questions, varying some data values and observing changes in final or intermediate results.

Spreadsheets, even if they are flexible and straightforward, are not the final answer. Some methods are too complicated to be efficiently coded in spreadsheets. In such cases, *statistical packages* can be used. The range of statistical packages is very wide. All traditional statistical packages, available for mainframes in the past, are now available on personal computers and workstations. There are many new packages available, heavily supported with graphics and attractive user interfaces. They range from simple and small to sophisticated and huge. It is not easy to decide which package to use in teaching, nor is my aim to recommend any specific package.

First of all, we must decide whether we want to simply teach statistical concepts or, at the same time to train students to use a certain statistical package as potential future users. In the latter case, the choice is influenced by availability of the package for future users, implementation of specialized methods important for the field of research, and so on.

If we want to select a statistical package for teaching, other aspects should be considered. Most statistical packages are powerful enough - if not too powerful - for teaching, because for teaching only a limited range of the available methods will be used. The selection should be based on ease of use, clarity of presentation of results and flexibility in data manipulation.

In the past, the burden of calculations was exchanged for the burden of complicated computer usage. Modern trends in computer usage are headed to "point-and-click" methods. It is certainly a great step towards "easy-to-use" but is also a step to "easy-to-forget". Maybe I am too traditional (or should I say old fashioned?) but I think that certain mixture of clicking and typing should be used. It is fine if a menu-driven or windows-like shell leads the student over pits and traps of computer usage, taking him to the core of selected application. But it is better to type in commands for statistical procedures since one has to think what one will type. I believe that things are remembered better if they are written down or typed in the computer than if they are just pointed or clicked on the screen.

In the area of presentation of results, two aspects should be considered: printed results and graphical presentation. Printed results are traditionally organised in lists and tables, but it is a nice feature if a log file for inspection of previous results is available. In some cases, results are kept in a specific object format that can be viewed at any time and even analysed by another method. Sometimes, results are presented in spreadsheet-like log file, parts of which can be copied to the data editor and reanalysed.

Graphics depend greatly upon hardware capabilities and will be discussed in the next section. In regard to data manipulation, many statistical packages incorporate spreadsheet-like data editors. The ability to apply mathematical and statistical functions to the data in a spreadsheet is a fine feature of such editors.

3. Statistical graphics

Graphical presentation of statistical data and results of statistical calculations is a very important part of statistical practice. Computers can greatly help the teaching of appropriate graphical presentation.

In the past, graphical capabilities of computers were limited to low-resolution printed bar charts, histograms or scatterplots. Modern spreadsheets and statistical packages are well equipped with a wide range of possible types of plots. A dataset can easily be graphed in many different ways and their appropriateness can be discussed in class. Such discussions are hardly possible without computers, because the effort for producing a lot of different graphs - some of which may be inappropriate - by hand is too large.

On the other hand, it is advisable to train students to be able to sketch graphs without the help of computer graphics. Only the usual types of graphs are implemented in computer programs, especially in spreadsheets, where more emphasis is given to presentation graphics than to analytical graphs. Some skills to construct uncommon or new graphs by hand should be developed.

4. Simulation

Statistical reasoning, based on concepts of probability, variation and uncertainty has to be supported by intuition or experience with data in different situations. Using computers one can repeat the same analysis on different, simulated, sets of data. Such simulated samples, taken from known distributions enable us to compare the results of an analysis of a sample with the known, exact state of the original distribution. Such repeated analysis can be conducted many times, showing to students the variety of possible results, giving them the opportunity to feel the probability of the expected outcome. In that way students can see the power and limits of applicability of statistical methods.

To support such simulations, generators of random numbers from different distributions should be implemented in spreadsheets and statistical packages. While some statistical analyses and concepts can be shown by simulated data with spreadsheets and general statistical packages, some should be shown by special programs, designed for statistical simulations.

Some very important statistical concepts, such as confidence intervals, sampling distribution or standard error, are difficult to understand for many non-mathematically inclined students. Specialized graphically supported *statistical experimental/simulation systems* (SE/SS) should be developed to show such concepts in an intuitive way. They should show not only standard results, but also the backstage performance of the statistical method. In the examples mentioned above, using repeated sampling from a simulated population, one can plot relevant sample indicators (sample itself, mean value, confidence interval...) for each sample and the sampling distribution at the same time. In that way, spectators can see a variety of samples and compare their properties with the known state of the population. It is important to show to students how representative or unrepresentative samples of the same size can be.

The purpose of SE/SS is to demonstrate difficult statistical concepts graphically. For non-mathematically inclined students formulas and proofs are not always enough. In some cases attractive, animated graphical simulation can help students to understand the underlying principles.

In that way, the effect of different sampling schemes and techniques can be shown. Among others, concepts of significance in hypothesis testing, problems of outliers in regression analysis and some aspects of analysis of variance are interesting for simulated presentation. Computer supported SE/SS can be incorporated into individual learning programmes and lab work, but they can also serve as illustrative tools in lectures.

5. Electronic textbooks

Electronic textbooks or programmed learning sequences are not limited to the field of statistics. A general trend is to implement them as a sort of self-explaining tool, suitable for individual learning at one's own pace. As in other disciplines, they are useful in statistics too. They provide an excellent mixture of classical textbook and self assessment. Compared to classical textbooks, they have the advantage of animated examples and graphs. Graphically supported simulations can be incorporated and explained by the surrounding text. Problems and questions are used as an instant test of understanding, leading the student to the next topic or guiding him back to reread some topics in the case of failure. Testing is not a one-way process and can serve as a guide for the author, indicating weak points in explanatory text in case of frequent student failures. A textbook can be improved with feedback between the author and the student.

Electronic textbooks have, at the moment, two great disadvantages compared to classical textbooks: they are rare and they cannot be studied elsewhere. The first problem can be solved by lecturers of statistics and the second by new technology. But even so, they are not practical enough to push paper based textbooks from student's hands.

I have to mention *multimedia*, a very promising field of application of computer technology to education. The multimedia approach is closely interconnected with electronic textbooks, unfortunately putting even greater demands to computer equipment, limiting the availability of such systems. While anyone can take a paper textbook to any place in the world, and be able to study it as well as anyone else, electronic textbooks are restricted in use.

6. Teaching with computers

Using computers influences the way of teaching. Some skills, traditionally taught, may not be needed in the future. Some calculation schemes, used in the past to ease the calculation process, are no longer used. Should they be abolished? Let me present some examples.

Calculation of the mean by the method of deviations from an arbitrary value is no longer needed if calculation is performed with computers. But it illustrates the concept of data transformation and connection between the statistics of the original and transformed data. One student asked me, why are we learning formulas for calculations of the arithmetic mean and variance from grouped data, if statistical packages calculate statistics from raw data. Maybe the computational reason vanished but as a concept they are still important.

Should we teach reading of values from probability the density tables, knowing that statistical packages report "exact" probabilities and significance levels? As it teaches the concept of a table and shows the connection with probability density functions, yes. And we can even do it better than in the past, if computer graphics are used to show what we actually read from tables.

We can compare such dilemmas with the role of the logarithm function in the past and at the present time. The logarithm was very important, and according to some indications, was invented to simplify multiplication to addition. It is nowadays no longer needed for that purpose, but is still very important as a concept of certain relations in mathematics and nature.

So we don't have less to teach, but I hope that we can teach better with computers. They can be used in individual lab work of students-mostly for calculation and graphing purposes. For me, more promising and interesting is the incorporation of computers, especially the demonstrations of simulations, into lectures. They should be used as animated transparencies for overhead projections. As is true with transparencies, we should pay attention to visibility of details. Characters should be of appropriate size to be read from back row of the classroom. Resolution for graphics should be reduced, plotting lines as bars, several pixels wide. Colours on the screen should be carefully selected if monochrome projection is used. Projectors should be luminous enough to present clear picture. If these recommendations are not met, the goal of the presentation will not be fulfilled, regardless of a carefully prepared spreadsheet calculation or excellent simulation.

7. Conclusion

Computers can be incorporated into the teaching of certain statistical topics.

Using computers, one can spare a lot of time otherwise devoted to calculations. We can show more extensive examples in lectures. Students can analyse larger samples, plot more graphs and get more experience with data. Different approaches to the same problem can be tested, graphs of different kind can be easily presented. As the burden of calculation or graphing can be left – to a certain extent – to computers, more time is left for the discussion and explanation of statistical concepts. Understanding of different statistical concepts can be enhanced by the use of computerized statistical experiments and simulations.

Computers can help us to teach statistics on some levels. Using computers in lectures is an additional effort for teachers. Computers should be used, if they are available, for tasks that cannot be efficiently shown on the blackboard.

Availability of computers is not a question to be neglected. In many schools around the world they are simply absent. Where they are available, they are frequently unable to run modern software. It is wise to keep in mind that teaching software must be able to run even on simple computers.

Computers, however capable they might be, will not push teachers from their desks in the near future. If a good teacher is a flame, boiling the statistics teaching stew, computers are spices: the stew is better if they are used with care, but is likely to be tasty without them too.

UNDERSTANDING PROBABILITY AND STATISTICAL INFERENCE THROUGH RESAMPLING

Clifford Konold
Scientific Reasoning Research Institute
Hasbrouck Laboratory, University of Massachusetts
Amherst, MA 01003, USA

1. Introduction

During the past four years, I have been developing *curricula* and computer software for teaching probability and data analysis at the introductory high-school and college level. The approach I've taken emphasizes the use of real data, where "telling a story" takes priority over testing hypotheses, and in which mathematical formalism is kept to a minimum (see Cobb, 1992; Scheaffer, 1990; Watkins *et al.*, 1992).

A major question I have considered is how probability ought to be integrated into this new data-rich curriculum. There are two major reasons for keeping probability in the data-analysis (or statistics) curriculum. First, at some point in the process of constructing theories that account for patterns in data, it is important that students consider alternative explanations. Among these is the possibility that some outcome of interest resulted from chance. Second, probability is an important concept in its own right (Falk and Konold, 1992). It comprises a world view and should not be viewed a necessary evil that must be faced if students are to understand statistical inference.

In deciding how to relate probability and data analysis, I have adopted an approach Julian Simon began advocating in the late 1960s. Simon describes his approach as having grown out of his frustration watching graduate students do silly things when trying to test a statistical hypothesis (Simon and Bruce, 1991). He began designing physical experiments from which he hoped they could build up sound probabilistic understandings. These eventually developed into a resampling approach that promised a more intuitive take on probability and data analysis, and which made the connections between the two fields more apparent. Of course, Simon didn't invent Monte Carlo methods, nor the randomization tests he would come to employ, but he was among the first to see their educational potential, and long before the computer was widely available.

Rather than elaborate Simon's argument here, I briefly describe two software tools we've developed, highlighting aspects that emphasize the relation between probability and data analysis. I also report some results from our primary test site, a high school in Holyoke, Massachusetts.

2. Modeling a problem with Prob Sim

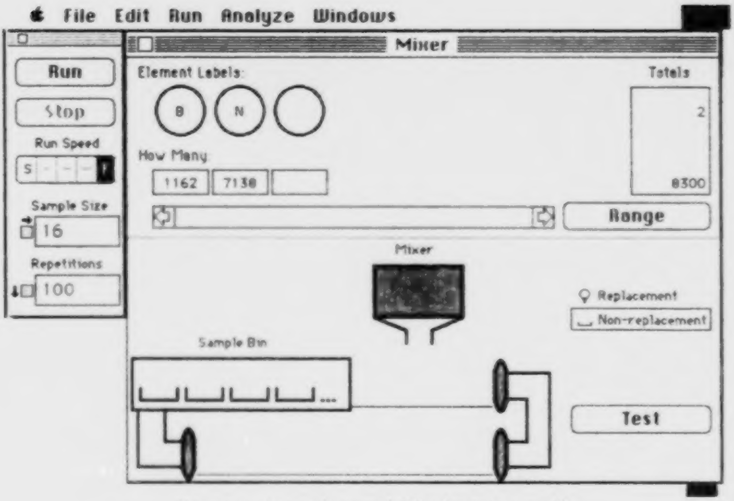
Most educational probability-simulation software comprise several ready-made models (e.g., coin model, die model). Students load the appropriate model, draw samples, and then see results displayed. The software thus offers empirical demonstrations of various facts and principles, such as the law of large numbers and the binomial distribution. The software we have designed, "Prob Sim", includes no ready-made models. Rather, the student must build the model, specifying the appropriate sampling procedure and analyses in order to estimate the probability of some event. The process of building a simulation model is at least as important as, if not more important than, drawing the appropriate conclusions from the results. To illustrate, I'll describe one of our activities entitled "LAPD" (see Konold, 1993, for another example).

Students begin by reading excerpts from *The New York Times* account (March 18, 1991) of the beating of Rodney King by officers of the Los Angeles Police Department. After reporting that "at least 15 officers in patrol cars converged on King", the article broaches one of the issues that made this incident explosive: "In what other police officers called a chance deployment, all the pursuing officers were white. The force, which numbers about 8,300, is 14% black".

Students are asked to build a model of this situation to estimate the probability of finding no blacks in a random sample of 16 officers. This problem has generated lively discussions in our test sites. When students care deeply about a problem, they are more willing to persist through difficulties. Moreover, they learn that application of probability theory is not limited to rolling dice and blindly drawing socks out of dresser drawers.

2.1 *Building a model*

I demonstrate below the various stages through which students progress in modeling this situation, illustrating the steps with screen shots from Prob Sim. In Fig. 1, a mixer has been filled according to the information provided in the article. There are a total of 8,300 elements, 1,162 of them labeled *B* (Black) and 7,138 labeled *N* (Not black). The non-replacement option has been selected to preclude the possibility of having the same element (officer) in a sample more than once.



Sample Size 16

Repetitions 100

Element Labels: B N

How Many 1162 7138

Totals 2 8300

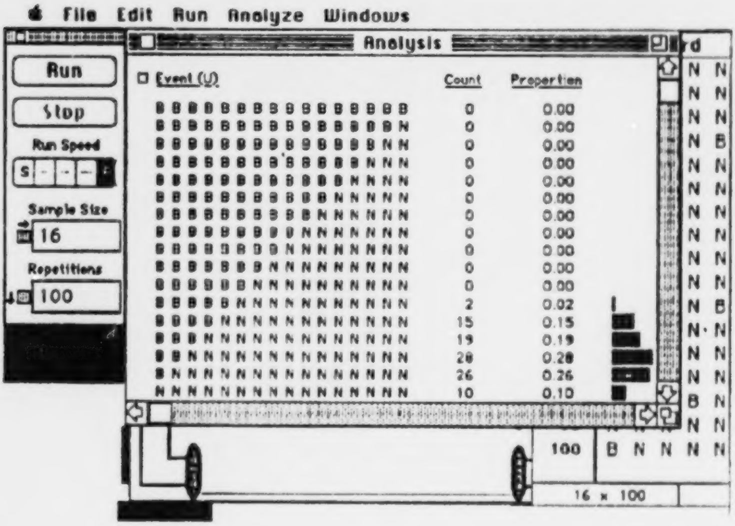
Figure 1. *A sampling model for the LAPD problem*

The Run controls on the far left shows the sample size set at 16, and number of repetitions set (somewhat arbitrarily) at 100. After the Run button is pressed, the computer draws 16 elements from the mixer without replacement, repeating this a total of 100 times. This is analogous to looking at 100 occasions when 16 randomly-selected officers appeared on a crime scene. The sampling process is animated in the lower part of the Mixer window. The results of each repetition are displayed in a Data Record window (not shown).

After the block of 100 repetitions has been drawn, results can be analyzed as shown in Fig. 2. The Analysis window in this case shows the number (and proportion) of occurrences of each of the 17 possible unordered outcomes. Ten percent of the samples had 0 *B*'s. Thus, a first estimate of the probability that a "chance deployment" would include no black police officers is .10.

2.2 *Repeating the experiment*

In many textbook examples of simulation, the experimental component ends here. But it is important not only to estimate the probability of some event, but to get some sense of the range (or variability) in that estimate given the number of repetitions. Indeed, the notion of chance is not apparent in a probability value per se, but in the variability of result over



Sample Size 16

Repetitions 100

☐ Event (U)	Count	Proportion
B B B B B B B B B B B B B B B B	0	0.00
B B B B B B B B B B B B B B B N	0	0.00
B B B B B B B B B B B B B B N N	0	0.00
B B B B B B B B B B B B B N N N	0	0.00
B B B B B B B B B B B B N N N N	0	0.00
B B B B B B B B B B B N N N N N	0	0.00
B B B B B B B B B B N N N N N N	0	0.00
B B B B B B B B B N N N N N N N	0	0.00
B B B B B B B B N N N N N N N N	0	0.00
B B B B B B B N N N N N N N N N	0	0.00
B B B B B B N N N N N N N N N N	0	0.00
B B B B B N N N N N N N N N N N	2	0.02
B B B B N N N N N N N N N N N N	15	0.15
B B B N N N N N N N N N N N N N	19	0.19
B B N N N N N N N N N N N N N N	28	0.28
B N N N N N N N N N N N N N N N	26	0.26
N N N N N N N N N N N N N N N N	10	0.10

Figure 2. *The analysis window showing the results of 100 repetitions*

replications of an experiment. To replicate an experiment in Prob Sim, the student has only to press the run button again. Another random sample is drawn, and the Analysis window updates to show the new results.

Students replicate the experiment about 10 times, plotting the results on a line plot which helps to emphasize the variability in the process. They then pool their results to come up with a final probability estimate. For the 10 results plotted in Fig. 3, the pooled estimate of P (0 blacks) is .109.

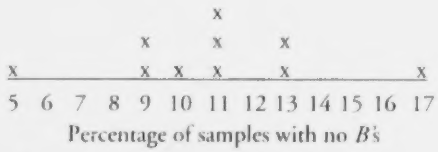


Figure 3. *Line plot of result of 10 replications*

In comparing results obtained in other groups, students discover that there is considerably less variability among the pooled estimates of the different groups than there is in the 10 estimates they got in their own group. Students thus see empirical demonstrations of the law of large numbers long before encountering a formal description. I've used such tasks in short two-week workshops in which I've never formally introduced the law of large numbers. According to performance on post-test items, these students show significant gains in understanding the effects of sample size on the variability of a sampling distribution.

When they can, students also work out theoretical probabilities. Through comparing these to their empirical values, they gain confidence in the simulation process as well as an appreciation for the power of mathematical theory and formalism.

2.3 *Questioning assumptions and making decisions*

In the process of model building, various assumptions are made. Students are asked, for example, to consider how the fact that police officers often arrive in pairs might affect the probability in question if partners tend to be of the same race. If they can, they build and run models consistent with these new assumptions and discuss the implications of their results. It is through addressing the assumptions incorporated into any model that students begin to understand what the process of modeling involves. In order to determine the predictive value of the simulation results, students must decide if and how various considerations affect the real situation. In some cases they can alter the model to take into account an additional factor (e.g., have police arrive on the scene in same-race pairs). If they can't do this, they can often predict the direction of biases introduced by their simplified model. Through this process of comparing a model to the target situation, students come to realize that though they can't avoid simplifying assumptions, the more aware they are of the limits of a particular model, the more informative the data from that model becomes.

At the end of the lab, students discuss the problem of deciding, based on their information, whether it was just chance that none of the responding officers were black. And, of course, the probability of the event is only one of the relevant pieces of information. More important are the subsequent testimony of the officers and the sensibility of any conspiracy theories that explain why black officers weren't on the scene. In other words, the probability of some event is never the only or last word in making some decision related to that event.

Such discussions seldom occur until instruction moves beyond the realm of coins and dice. Prob Sim is an important tool in supporting discussion of this type, because it can be used to model engaging problems too complex for students to tackle formally. Prob Sim's simplicity, and the speed with which sufficient data can be generated, gives students the time to design a sampling model, collect adequate data, draw conclusions, and discuss implications of the findings. In the next section, I show how we build on ideas introduced in modeling probabilistic situations when we move on to data analysis.

3. Data analysis using DataScope

We designed a data-analysis program called "DataScope" for use in introductory courses stressing exploratory data analysis in which students work with real data. Our primary objective is for students to learn basic data-analysis techniques for exploring relationships among various variables. By using multiple-variable data sets, we give students the opportunity to explore questions of particular interest to them. DataScope encourages students to make initial judgments of relationship by visually comparing plots. This is demonstrated below using data obtained from a questionnaire administered to 84 students in two high schools in western Massachusetts: Amherst Regional and Holyoke High. Amherst is a small college town, while Holyoke is a larger, industrial city. The information collected on each student included gender, age, birth order, family size, marital status of parents, religious activity, rating of school performance, educational level of parents, curfew times, working hours and wages, and time spent on homework. Students spend about two weeks exploring various questions in this data set, among them the question of whether holding an after-school job adversely affects school performance.

DataScope encourages exploration by allowing students to form subgroups of some variable based on the values of some other, presumably related, variable. For example, Fig. 4 shows the box plots for hours of homework (HWHRS) and hours spent working (JOBHRS) for Amherst and Holyoke students (in the case of JOBHRS at Amherst, the median is the same as the 1*st* quartile, as indicated by the double-thick line). The number of cases in each box plot is displayed to the right of the plot. This is important to include as students will frequently draw conclusions without considering the number of cases represented by each box plot. The plots below suggest that Amherst students spend more time on homework than working a part-time job, and that the opposite is true of Holyoke students. This trend is consistent with commonly-held stereotypes of the two towns.

It is tempting to conclude from Fig. 4 that those with jobs spend less time on homework than those without jobs. However, when HWHRS are "grouped" by the categorical variable JOB, students discover the reverse appears to be true – students with a job studied an average of three hours-per-week more than those without a job.

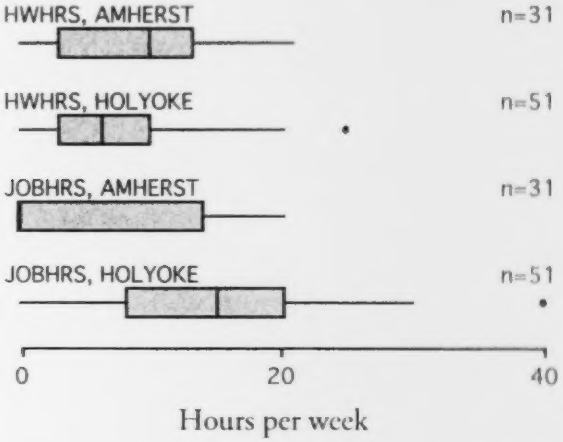


Figure 4. *Box plots of homework hours and job hours for Amherst and Holyoke students*

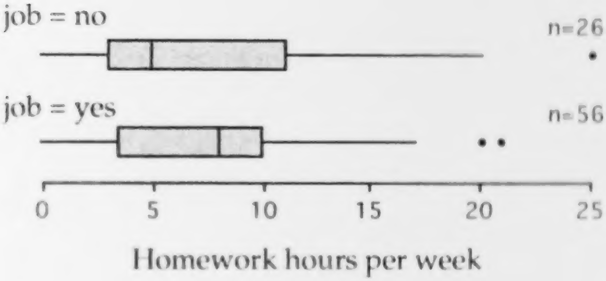


Figure 5. *Box plots of homework hours for student with and without jobs*

At this point, students reconsider their expectations and develop theories that might explain these data. One possibility that those who work also study more because they have learned to effectively manage their time. Or maybe the students with jobs are older and have more homework assigned. Some of these explanations can be investigated by looking at other variables in the data set. However, among the explanations to consider is the possibility that the difference was due to the "luck of the draw" – that with a better or larger sample, no difference would be found. In my experience, students do not spontaneously raise this possibility, even when they have been using resampling techniques, as described above to estimate probabilities. They will judge differences between two medians as important in one case, and unimportant in another, apparently making the judgment based on the distance between the two medians, as it appears on the computer screen. They do not spontaneously evaluate this difference with respect to the variability (i.e., to the IQRs as shown in the box plots). Therefore, before showing them how we might determine the probability of a difference occurring by chance, I typically must remind them that this is a possibility.

3.1 *Randomization tests in DataScope*

Below is a demonstration of how DataScope is used to estimate the one-tailed probability of observing a difference at least as large as the one observed in the sample above. The method is based on the randomization procedure as originally developed by Fisher (see Barbella *et al.*, 1990). This involves randomly reordering one of the variables (without replacement) to estimate the probability of the observed difference under the null hypothesis. With DataScope, the student can first do this "manually", to develop a sense of what the computer is doing. A "reorder" command randomly reorders the values of one of the selected variables (in this case, the values for HWHRS). That is, the values of HWHRS are randomly reassigned to cases. Once reassigned, the variable name appears in the data table with the symbol ® on both sides to remind the student that the column has been randomized (another command will restore the original order). A box plot of this randomly ordered variable, again grouped by the job variable, can be viewed. Given that the values of HWHRS have been randomly assigned, any difference between the medians of the two groups (with or without jobs) is due to chance. In the example shown in Fig. 6, this difference is -1 (sub-tracting the median of the upper plot from the median of the lower plot).

Once students understand the procedure, the computer can be instructed to repeatedly reorder the variable and compute the difference between medians of the job and no-job groups, recording these in a new data table as illustrated in Fig. 7. In this case the computer has been instructed to draw 100 random samples. The first few differences that were obtained are shown in Fig. 7 in the "Resampling" data table. The first value is the observed difference, -3. In the background you can see the variables selected (V1 and G1) in the primary data table. The values in HWHRS are being randomly reordered.

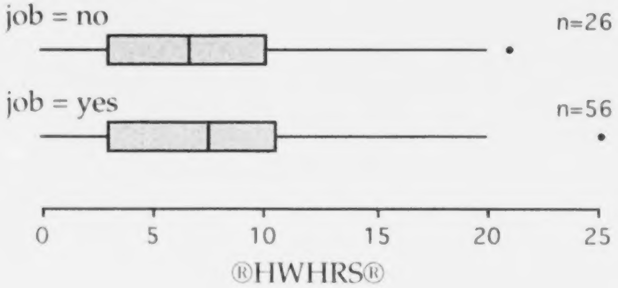
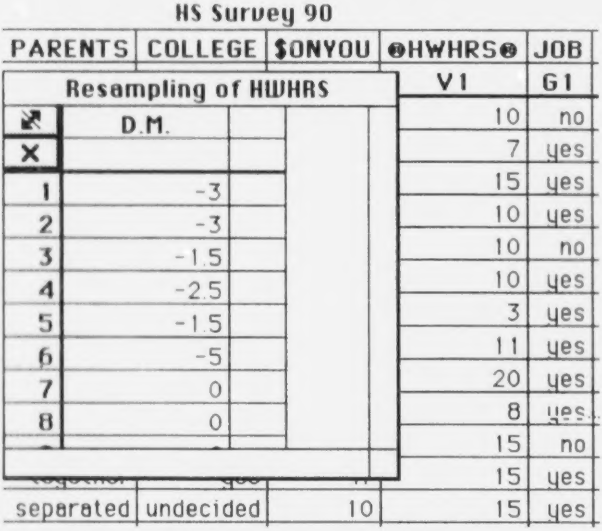


Figure 6. *Homework hours randomly assigned to job and no job groups*

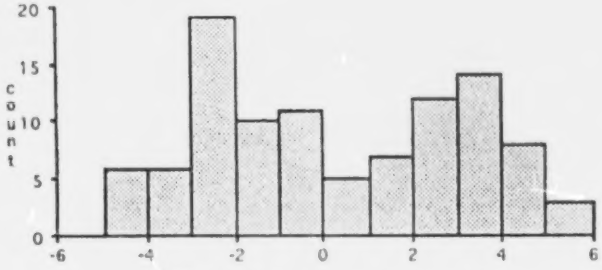


	D.M.			
1	-3			
2	-3			
3	-1.5			
4	-2.5			
5	-1.5			
6	-5			
7	0			
8	0			

PARENTS	COLLEGE	$ONYOU	®HWHRS®	JOB
			V1	G1
			10	no
			7	yes
			15	yes
			10	yes
			10	no
			10	yes
			3	yes
			11	yes
			20	yes
			8	yes
			15	no
			15	yes
separated	undecided	10	15	yes

Figure 7. *Table of differences between median HWHRS for job and no-job groups in successive resampling runs.*

After the 100 random differences have been obtained, the results can be displayed in a histogram as shown in Fig. 8. In this instance, 25 of the samples had a difference at least as large as -3. Thus, an estimate of the one-tailed *p* value is .25. Additional repetitions could be conducted for a more precise estimate. This same procedure is used in DataScope to test the statistical significance of a value of *r*, and of frequency counts in a 2 X 2 table.



Difference between medians

Figure 8. *Sampling histogram displaying results of 100 resamplings*

4. Educational outcomes

I have found, as have Simon and Bruce (1991), that students are enthusiastic about probability and statistical inference when approached through resampling. But, do students using this approach learn more than they do in a traditional course? Simon *et al.* (1976) compared student performance in courses taught using resampling vs. conventional methods. Given problems that could be solved using methods taught in either course, students in courses using the resampling approach consistently outscored students using the traditional approach.

Many or most students who take an introductory course will never need to conduct a statistical test or determine a probability precisely. They do, however, as members of a complex and increasingly technological society, require a basic understanding of uncertainty and the savvy to evaluate "research" claims in the mass media. Accordingly, though students should be able to solve problems using methods they have been taught, it is even more important that they understand basic concepts which underlie these methods. Konold and Garfield (1993) have developed items to assess understanding of these basic ideas. Below is one of our problems, adapted from an item in Falk (1993, p. 111).

> The Springfield Meteorological Center wanted to determine the accuracy of their weather forecasts. They searched their records for those days when the forecaster had reported a 70% chance of rain. They compared these forecasts to records of whether or not it actually rained on those particular days.

The forecast of 70% chance of rain can be considered *very* accurate if it rained on:

a. 95% - 100% of those days.
b. 85% - 94% of those days.
c. 75% - 84% of those days.
d. 65% - 74% of those days.
e. 55% - 64% of those days.

This problem was designed to assess whether a student understands that a probability is a *quantitative* measure of uncertainty (or frequency of occurrence). It is difficult to imagine how, without this understanding, a student could correctly interpret a p value, e.g., that if we estimate the probability in the LAPD problem as .10 this means that roughly 10% of the time when a random sample of 16 officers is selected, there will be no blacks among them.

Fig. 9 shows the frequency of responses to the various options in the weather problem by 199 students we have administered this item to before instruction. Forty-three students (32%) selected the correct range 65-74, which includes within it the normative value of 70%. The majority of students (36%), selected the highest range, suggesting that they expect rain to occur nearly all of the time when it has been forecast with a 70% probability. This answer may be based on an approach to uncertainty that is prevalent among adults across a range of problems. Because it is based on the belief that the objective in probabilistic situations is to predict the *outcome* of a *single* trial, I have referred to this as the "outcome approach" (Konold, 1989).

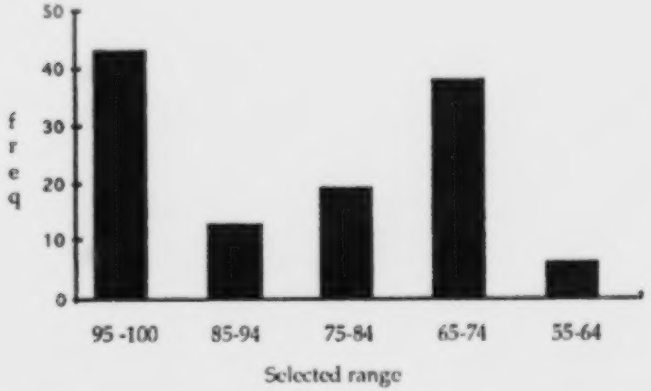


Figure 9. *Frequency of responses before instruction of 199 students to the weather problem*

One of my instructional objectives, therefore, is for students to realize that a probability value typically tells us little or nothing about results in the short run, but a great deal about results in the long run. Fig. 10 compares the results on this same problem before (black) and after (gray) instruction. Correct responses increased only 6% with instruction. The results are similar across the majority of our assessment items. At a deeper level, many students after instruction using resampling appear unaware of the fundamental nature of probability and data analysis.

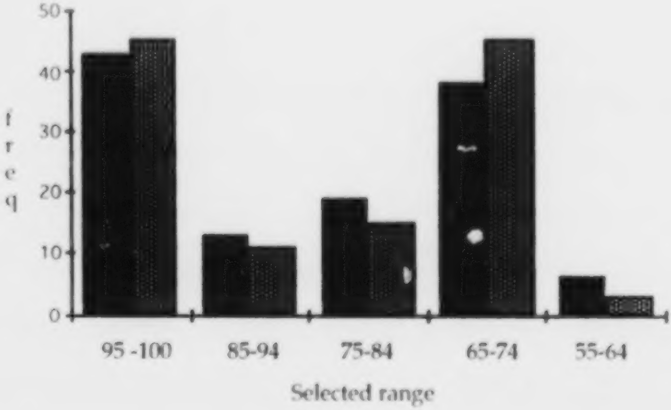


Figure 10. *Frequency of responses before (black) and after (gray) instruction of 199 students to the weather problem.*

I remain optimistic about the instructional benefits of the resampling approach. I have used it in individual tutoring sessions with students at a variety of levels who express momentary insight into the logic it employs. If the benefits have not been apparent in our larger implementations, it may be due to the fact that we have a broad range of objectives and have not devoted enough time to any particular concept. Too, it is easy to underestimate the conceptual difficulty of probability and chance. The fact that probability was a late bloomer, historically speaking, should help alert us to the conceptual complexities which are belied by the simple formalisms. Perhaps looking at development over a single course is too small a unit of analysis in the case of probability, and we should be thinking about series of courses over which we can expect to effect and observe conceptual change.

Bibliography

Barbella P., Denby L. and Landwehr J.M. (1990), Beyond exploratory data analysis: The randomization test, *Mathematics Teacher*, 83, pp. 144-149.

Cobb G.W. (1992), Teaching statistics: More data, less lecturing, in L. Steen (ed.), *Heeding the Call for Change*, (MMA Notes # 22, pp. 3-43), Mathematical Association of America.

Falk R. (1993), *Understanding probability and statistics: Problems involving fundamental concepts*, Wellesley, MA.: AK Peters (to appear).

Falk R. and Konold C. (1992), The psychology of learning probability, in F.S. Gordon and S.P. Gordon (eds.), *Statistics for the twenty-first century* (MMA Notes # 26, pp. 151-164), Mathematical Association of America.

Konold C. (1993), Teaching probability through modeling real problems, *Mathematics Teacher* (to appear).

Konold C. (1989), Informal conceptions of probability, *Cognition and Instruction*, 6, pp. 59-98.

Konold C. and Garfield J. (1993), *Statistical Reasoning Assessment. Part 1: Intuitive Thinking*, Scientific Reasoning Research Institute, University of Massachusetts, Amherst.

Scheaffer R. L. (1990), Toward a more quantitatively literate citizenry, *The American Statistician*, 44(1), pp. 2-3.

Simon J.L. and Bruce P. (1991), Resampling: A tool for everyday statistical work, *Chance: New Directions for Statistics and Computing*, 4(1), pp. 22-32.

Simon J.T., Atkinson D.T. and Shevokas C. (1976), Probability and statistics: Experimental results of a radically different teaching method, *American Mathematical Monthly*, 83, pp. 733-739.

Watkins A., Burrill G., Landwehr J. and Schaeffer R. (1992), Remedial Statistics?: The implications for colleges of the changing secondary school curriculum, in F.S. Gordon and S.P. Gordon (eds.), *Statistics for the twenty-first century* (MMA Notes # 26, pp 45-55), Mathematical Association of America.

Note

The curriculum materials and research described in this article were supported by a grant from the National Science Foundation (Grant #MDR-8954626). The opinions expressed here are the author's and not necessarily those of the Foundation.

THE USES OF VIDEO IN TEACHING STATISTICS

David S. Moore
Department of Statistics
Purdue University, West Lafayette, IN 47907-1399, USA

1. Introduction

Video, in the form of broadcast television, is the most popular medium for entertainment and news in the developed world. A television set is one of the first substantial purchases made by households in developing areas as their wealth increases. These phenomena testify to the power of video to hold attention, a power which can also be applied to formal teaching. Video can be used for teaching in several settings: learning at a distance for geographically scattered students, as a supplement in traditional classroom settings, and as a component of new technological learning systems. In each case, wise use of video requires an understanding of both the strengths and weaknesses of the medium. This paper will review these strengths and weaknesses, drawing both on practical experience and on cognitive research. We will then suggest appropriate uses of video for teaching statistics in the three settings just mentioned.

2. Strengths of video in teaching

One of the most obvious advantages of broadcast television is *reach*, the ability to convey information at a distance. The power of television compared with radio lies in the *visual* aspect of the information conveyed. It is here that the greatest advantages of video reside, advantages that are unrelated to geographic reach.

Most essentially, *video compresses time and space by making distant or past events present in visual form.* Students of statistics can visit clinical trials, government statistical offices, and factories to see and hear medical researchers, government officials, and quality engineers describe and apply statistical techniques. Direct viewing replaces the attempt to imagine a setting described by the instructor. Video compresses time and space in other ways as well. Slow motion and time-lapse photography expand the scale of time over which phenomena are visible to us, and views recorded from microscopes and telescopes expand the scale of space that is accessible. Students can observe the growth of crops or tumors (time lapse), the etching of semiconductor chips (microscopy), and interviews with sample survey participants. Applications of statistics become both more understandable and more interesting.

Even without these explicit extensions of scale, carefully edited video material compresses space and time: we move quickly from place to place and across time to concentrate on essential aspects of the phenomena presented. We can observe in quick succession the planning of a sample survey, the training of interviewers, excerpts from interviews with respondents, followup, data coding and analysis, and presentation of the results. For clear, concentrated presentation of a message, video is often more effective than personally viewing the phenomenon. From the early days of film this power has been used to persuade, both for propaganda and for advertising. We can also use it to instruct.

A second strength of video is the ability to present *dynamic graphics.* Movement and color enhance the effectiveness of many data graphics. Contemporary computer graphics recognizes this, but even traditional graphs are often easier for students to grasp when presented dynamically with narrative comment on each feature. Consider Minard's now-classic graph of the fate of Napoleon's army in the Russian campaign of 1812 and 1813 (Tufte, 1983, p. 41). This may be, as Tufte says, "the best statistical graphic ever drawn", but it records six variables and requires very careful study to understand. A dynamic version that unfolds the campaign month by month, ending with Minard's static graphic, is much easier to follow.

Video offers to statistics teaching standard graphics of a quality beyond the level (for now) of most student computer systems. More important in the long run, video can present custom graphics such as a dynamic version of Minard's classic. Some statistical concepts, such as sampling distributions, are inherently dynamic: "How does this statistic behave in many repeated samples?" A custom dynamic graphic that represents the drawing of repeated samples and pictorially builds up the sampling distribution is a natural teaching tool.

Finally, we must note the subliminal effects of video as well as its more obvious advantages. *Video communicates more through feelings and emotions than through linear, rational thought.* This fact, suspected by many critics of the cultural effects of television, is substantiated by psychological studies (Howe, 1983). Video changes attitudes more effectively than it conveys detailed information. Awareness of this aspect of video is essential to informed use of video in teaching. We show video clips of statistical applications not merely to set the stage for exposition of statistical methods, but to change students' attitudes towards statistics.

3. Weaknesses of video in teaching

Video as a tool for teaching has striking strengths, but also some striking weaknesses. First, it is *very expensive to produce.* The real-world applications, shot on location and carefully edited, that are video's most valuable contribution to statistics teaching, are particularly expensive. Production costs for U.S. commercial television exceed $ 500,000 per half-hour program, and many fast action spectacular advertising spots cost that much for 30 seconds. "Talking heads" are much less expensive, but take little advantage of video's greatest strengths.

Technological progress will bring down the cost of video to some extent. Video is now an analog medium, so that editing is similar to frame-by-frame editing of film and transferring computer-generated graphics to video requires special equipment. Video will soon be digitized, so that editing and graphics will be both faster and cheaper. However, location shooting will still require sending camera crews and their equipment to the action. This will always be expensive.

Also, *video is unsuited for detailed explanations and carefully worked examples.* The expense and rapid pace of video, as well as its multichannel nature, make these staples of teaching more suitable for text presentation. It is of course possible, and in some cases even desirable, to simply present a teacher teaching. But a talking head takes little advantage of video's specific strengths. Note that presentation of real-world applications is not an exception to the principle that video is unsuited for detailed explanations. Students who approach video applications as if they were text-style expositions become frustrated: they can't take notes because the presentation moves quickly and has several simultaneous levels which they cannot satisfactorily reduce to a single linear track. Video *shows* rather than tells. It is not an ideal vehicle for exposition.

Most important, *video in its current forms leaves its viewers passive.* They are simply spectators, not actively engaged in their own learning. The passive nature of video viewing is no doubt closely connected with its effects on emotions and attitudes. Current thinking about education emphasizes that students must be actively engaged in their learning, that passive reception of a message from even a live teacher has little effect (Garfield, 1993). Video may change attitudes, but without interaction it is unlikely to be effective in helping students understand concepts, make judgments, and apply skills to complex problems.

4. Video for distant learning

The reach of television makes it attractive for distant learning. The spread of video recorders (now present in over 70% of U.S. households, for example) aids distant learning by freeing the student from the broadcast hour and allowing repeated viewing. The visual aspects of television are essential for disciplines like statistics; radio does not broadcast equations effectively.

Distant learning has obvious disadvantages compared with contact with a live teacher. It is therefore most successful in specific circumstances. A relatively closed or expensive higher education system leaves unfilled demand that distant learning can satisfy. The somewhat closed nature of the older British university system is often cited as a reason for the initial success of the Open University. Geographical distance in areas that are either thinly populated or have few institutions of higher education can create demand for distant learning, as in the case of Open Learning Australia. Specialized training for already-employed students lies behind the widespread televising of post-graduate engineering courses in the U.S., most notably by the National Technological University. In all cases, students are most attracted when distant learning leads to a valued educational credential.

Advanced students will tolerate the talking heads of experts in specialized technical subjects, and have learned to interact with lectures on their own, guided only by problem assignments. Distant learning for advanced students is a setting in which the reach of television cancels its many weaknesses. In general, distant learning requires motivation and maturity on the part of students in order to compensate for reduced interaction with teachers. It is therefore most effective for older or more advanced students. The Open University generally requires students to be at least 21 years old.

Effective distant learning programs understand the weaknesses of video. Video is only a part, and often not the most important part, of a full program. Text material is essential, as is some arrangement for interaction with students via mail, telephone, or modem. Anthony Bates (Bates, 1983) notes that the Open University, a pioneer in distant learning, has gradually reduced the role of video both because of its expense and because of a growing understanding of the best uses of video. His essay is a most valuable account of the lessons of practical experience with distant learning. While distant learning can be effective, and is essential in special circumstances, it is generally "second best" when compared with live instruction.

5. Video in the classroom

All agree, for reasons outlined in discussing the weaknesses of video, that *video is not suitable as the sole medium of instruction.* On the other hand, the strengths of video are impressive. What is more, there is evidence that different people learn best via different media, and that video is particularly effective for people with high visual ability (Pezdek *et al.*, 1987). There are therefore strong arguments for including some video in a mixture of media for instruction.

The first principle for effective classroom use of video is obvious: *take advantage of video's strengths and avoid its weaknesses.* Video in the classroom ought not to be considered as a substitute for the teacher. In particular, basic exposition is best done by text and teachers, not by video. Short video excerpts consisting of dynamic graphics or of short documentaries describing statistical applications are very helpful. Documentary segments can motivate a new statistical topic. They can also provide a problem setting which we ask students to discuss with an eye to applying the statistical tools they are learning. I elsewhere (Moore, 1993) offer some specific examples of effective uses of excerpts from the U.S. Corporation for Public Broadcasting telecourse *Against All Odds: Inside Statistics.*

Teachers should remember the great effectiveness of video in changing attitudes. Regular showing of short documentaries illustrating statistical applications is justified by its effects on students' attitude toward statistics quite aside from any specific pedagogical purpose served by each individual excerpt. We need not inquire too closely about the specific technical content conveyed by a five minute visit to a clinical trial. It is enough to know that weekly viewing of such segments persuades students that statistics is useful and important much more convincingly than simply telling them that this is so.

A second principle for classroom video use is to *create interaction.* Ask students to discuss what they have seen. If the video segment is more than a few minutes long, stop the tape to ask questions, especially questions that force students to anticipate what comes next in the video. The teacher can compensate for the most serious weakness of video, the passive role in which it places students.

6. The future of video: multimedia tools for teaching

One clear conclusion from both research and experience in educational uses of video is that video is most effective in combination with other media. We can then tailor the use of video to take advantage of its strengths while the other media compensate for its weaknesses. Video has, I have argued, a potentially valuable role in combination with a printed text and interaction with a teacher. Now, however, new technology allows video to be integrated with text, sound, animated graphics, cartoon-style animation, and computing in an individual student workstation. This is a *multimedia* system. The student interacts with the system via keyboard, mouse, and perhaps voice. Advocates argue (not for the first time) that technology will at last supplant live teachers for much instruction. Ambron and Hooper (1990), Grabinger *et al.* (1990), and Steinberg (1991) discuss various aspects of multimedia technology and its use in education.

At the heart of multimedia is the fact that video, now largely analog, can be digitized. There is then no strong distinction between video and computing. Software can compress and store video images and play them in a window on the computer screen. Compression algorithms and storage technologies are important: full-motion video requires 30 frames per second, each frame a full picture's worth of bits. Current multimedia systems store video, still images, text, and software together on a compact disc similar to those used for music. A single CD can hold (in one current system) 650,000 pages of text, five hours of FM stereo sound, one hour of full-creen full-motion video, four hours of quarter-screen full-motion video, or any combination of these.

Multimedia has two clear strengths for education. The first is often called *hypermedia*: rather than a single linear track as in a book or movie, hypermedia offers multiple interconnected paths manipulated by the learner. The screen usually contains several windows, which can present different kinds of information or activities. Students can choose among several levels of exposition, can review a concept by clicking on a keyword, can go at any time to a window in which a statistical software package is operating.

Multimedia's second strength is that it expects *interaction* by the learner. The hypermedia idea requires choice, not simply reading on in a text. In particular, students can manipulate and interact with video rather than merely view it. For example, the learner can choose slow motion, or can freeze a frame and make measurements from the video to produce data for analysis. (Unlike tape, a CD does not wear when a frame is frozen). Text in another window can guide the viewing and interaction by questions or instructions. The data can be entered into a statistical software package (also in a window) for analysis.

Multimedia systems do what is needed to make video at least minimally effective for teaching: combine it with other media and require interaction. The technology does to some extent what a teacher should do when using video. Of course, it isn't easy to simulate a good teacher. The design of effective multimedia systems for teaching a subject like statistics presents many challenges. Some of the most impressive examples of multimedia instruction allow students to "explore a world". The learner can, for example, browse in an art gallery (represented on the screen) and click on any work presented to learn more about the era, the artist, or the specific work. Browsing and exploring are probably not sufficiently structured for beginning instruction in statistics. The many choices offered by hypermedia can overwhelm beginners, so that controlling complexity and finding a balance between learner choice and imposed structure are essential and difficult tasks.

Although multimedia technology requires interaction by the learner and thus attacks the most serious weakness of video instruction, it is not at all clear that the available interaction is sufficient for effective learning of higher-order skills such as solving multi-step, vaguely-posed problems and communicating the results. The computer finds it difficult to ask or answer "What do you mean?" "Why do you say that?" "Can you tell me more about the purpose of the study?" Yet such dialog is common in conversations between statisticians and clients and between good students and good teachers. Conceptual learning may still require conversation with a human teacher. Group problem-solving and clear communication of findings – essential in statistical practice and central to reform of statistics teaching – also require interaction with other people.

7. Conclusion

Video is a technology of great persuasive power that has a definite place in teaching at all levels. The reach of television is important in distant learning, and the impact of video is valuable both in traditional classrooms and as part of multimedia systems. New technology that incorporates video is a yet more powerful tool, which we should develop and employ with enthusiasm. But it is likely that, like video, newer technologies will not entirely replace human teachers. Just as video has strengths and weaknesses, we teachers have more-effective and less effective roles. Careful thinking about teaching will increasingly ask what roles each component – including the human teacher – can best play. In this paper I have tried to say what is the proper place of video. Similar information is available or being discovered for animation, computing (e.g., Biehler, 1993), sound, and text. We ought to aim to replace human efforts with technology where technology is effective, and to reserve valuable human time for roles that humans do best.

Bibliography

Ambron S. and Hooper K. (1990), *Learning with Interactive Multimedia: Developing and Using Multimedia Tools in Education*, Microsoft Press, Redmond, Washington.

Bates A.W. (1983), Adult learning from educational television: The Open University experience, in M.J.A. Howe (ed.), *Learning from Television: Psychological and Educational Research*, Academic Press, London, pp. 57-77.

Biehler R. (1993), Software tools and mathematics education: The case of statistics, to appear in W. Dörfler, C. Keitel and K. Ruthven (eds.), *Learning from Computers: Mathematics Education and Technology*, Springer, Berlin.

Garfield J. (1993), How students learn statistics, *International Statistical Review* (to appear).

Grabinger R.S., Wilson B. W. and Jonassen D.H. (1990), *Building Expert Systems in Training and Education*, Praeger, New York.

Howe M.J.A. ed. (1983), *Learning from Television: Psychological and Educational Research*, Academic Press, London.

Moore D.S. (1993), The place of video in new styles of teaching and learning statistics, *The American Statistician*, 47, pp. 172-176.

Pezdek K., Simon S., Stoeckert J. and Kiely J. (1987), Individual differences in television comprehension, *Memory and Cognition*, 15, pp. 428-435.

Steinberg E.R. (1991), *Teaching Computers to Teach*, second edition, Erlbaum, Hillside, New Jersey.

Tufte E.R. (1983), *The Visual Display of Quantitative Information*, Graphics Press, Chesire, Connecticut.

EDUCATIONAL HYPERTEXT IN THE TEACHING OF STATISTICS: A POSSIBLE USE FOR AN INSTRUCTION MODULE FOR THE ANALYSIS OF THE RELATIONS BETWEEN TWO VARIABLES

Furio Camillo[1]
Dipartimento di Scienze Statistiche, Università di Bologna
Via Belle Arti 41, 40126 Bologna, Italy
Marco Pedroni[2]
Logòn s.r.l., Via Formignana 1a, 44100 Ferrara, Italy

1. The hypertext as guide and conditioning: why use it in teaching?

The common definition of a hypertext is that it is a non-linear text or a collection of information arranged on a computer in such a way that access to it need not necessarily be of a sequential type. The information may actually be contained in texts or it may be in the form of graphic or photographic images, animations, films or sound recordings. Using special software the information is then included in files, whose consultation/reading is tied to graphic objects such as buttons, fields, windows, icons, etc. that allow interaction with other data or information on file (Corcione and Di Tonto, 1990; Barret, 1989).

Another definition of hypertext, which is suited to its use in an educational context, is that it is essentially a virtual text, i.e. a text in which the information does not have its own intrinsic ordered structure. In a hypertext, the information has only a virtual order, i.e. an order given by the possibility that the reader has to navigate from one anchor to another, passing through the nodes provided (the key pieces provided in the base structure). It is said, therefore, that particularly in the field of education the hypertext is a class of possible texts, generally very wide, but strictly determined, one of the members of which only becomes actual at the moment of reading. More precisely, each member of this class of texts becomes actual only through the operation of reading/consultation on the part of the user. Otherwise the member remains existent only virtually. It appears clear, moreover, that this virtual form of the organised information can exist only on the computer. Every print-out and every reading must necessarily be actualised, that is made concrete, by a particular request or instance (Antinucci, 1991).

---

[1] The paper is the result of the collaborative research of the two authors. F. Camillo wrote this paper and helped to construct the demonstration hypertext *BivarZero.*

[2] M. Pedroni contributed to the writing of section 3 and composed *BivarZero* in Toolbook.

The cognitive instance by which the user moves with respect to the hypertext thus represents the determining cause of the system of options and paths that this same user will follow from time to time. The choice of one path rather than another thus depends on the modes of perception of external stimuli (perturbations) that the subject normally adopts every day in the three phases of information processing: perception, signification and memorisation (Rondot, 1993).

In educational hypertexts, as in every other didactic tool, perception and memorisation are influenced by the sensibility and experience of the author. The author, however, inevitably loses a part of his or her authority over the hypertextual work, to cede it to the reader/user. It is the latter, in fact, who constructs step by step the fruition process, and hence the learning steps (Antinucci, 1991). Finally, in learning from a hypertext, signification is generally stimulated and directed by the author through the definition of the architecture of the whole hypertext, i.e. by the structure of nodes and links laid down *a priori*.

With regard to the possible didactic functionality of hypertexts, it has recently been emphasised that they must not be seen only as a new form of author system, able perhaps to produce courseware of a new type or global self-teaching environments. Rather, they must be considered also as an important tool for the arrangement and subsequent fruition of didactic materials and contents within a broader, more complete educational process (Varisco, 1991).

In company and professional environments, in particular, there are already numerous examples of hypertexts acting as didactic tools for reviewing or arranging knowledge, especially in the fields of technology and engineering, history and literature, architecture and design. In these cases, the hypertext allows the user to review and to investigate a vast collection of cognitive materials, linked together at the level of the company or professional courseware, which is constructed by the user himself (through the paths and options chosen) according to the particular needs of the moment or the particular problem that has arisen in the everyday productive sphere.

2. Why use the hypertext in teaching statistics? How to interpret the guide and the conditioning

The general background to which we have referred in our introduction to the hypertext is fairly traditional and commonly accepted among experts and scholars in the field. The use of hypertexts in the teaching of statistics, which is the subject of the remainder of this paper, is thus not of itself innovatory, except in terms of the context, the motivations and the purposes that have encouraged and directed our research.

One of the many definitions of Statistics says that it is "... the quantitative analysis of the observations of any phenomenon subject to variation, that is of a variable; objects for analysis are therefore all those phenomena which are subject to change..." (Naddeo, 1984).

Statistics is thus a vast subject, multiform in its subjects and open to innumerable approaches in the resolution of problems concerning the analysis of observations. Moreover, given the proliferation at every level of data and information collection processes, Statistics is used increasingly in fields that do not strictly belong to the statistician, and where it is increasingly considered as a tool for measuring and analysing the relations that exist between phenomena that are very much of a day-to-day operational nature.

In the light of these premises and of the above interpretation of the hypertext as a text that exists only in virtual form, it would seem to us that hypertexts may aid the learning and use of statistics by means of three fundamental kinds of intervention:

1) systematising, in a single large connected system, the strategies of analysis, research methodologies and techniques that make up Statistics;
2) providing the user of Statistics with a tool which directs him to the solutions of specific problems, but which allows him at any moment to construct his own map of statistical knowledge prior to the comprehension of the strategy itself (e.g., the necessary mathematical formalisations of the problems);
3) constructing an informatic environment that makes it possible to provoke a set of important cognitive perturbations in the learner. In particular, this means provoking perturbations that induce the user to turn directly towards the solution of the problems of analysis and which, through graphic and sound interaction, stimulate his intuitive ability to verify the solution identified, even before the learner possesses the necessary formal tools.

The links between the various parts of Statistics and between these and other disciplines could also represent a system of links of a sort of hypertextual *opera omnia* on Statistics about which it is difficult even to make hypotheses at the present stage.

There is no doubt that the hypertextual tool permits us to explain a complex multilevel, multipurpose and multidimensional structure like that which exists in the field of statistics education. Furthermore, for didactic purposes, it allows us to move users through the various areas of statistical knowledge by conditioning them to intuitive stimuli and directing them in their choice of paths. Particular attention must be paid, however, when planning a hypertextual statistics module. Navigation between the nodes and the anchor must be easy and, above all, there must be no danger of losing one's way in the intermediate cognitive charts with respect to a specific targeted path. Therefore, it is necessary to provide an easy return to the main path (relating to the specific statistical problem) from those propaedeutic formalisation/clarification charts relating to any problem the user may have opened up.

3. The module constructed (*BivarZero*): an experiment for concrete problems

The hypertext that we have constructed up to now represents only a first experiment. The *BivarZero* module concerns the problem of measuring and analysing the relation between two variables[3]. The spirit with which it was designed corresponds to what was said above in points 2) and 3), namely to the criteria of the strong applicability/operativeness of the solutions and to the intuitive characteristics of the user's choice of the best statistical solutions. Thus, the module created may be considered as a prototype for a larger courseware aimed at users interested in the application and operation of Statistics. The environment within which it was constructed is ToolBook 1.5 for Windows 3.0 for Personal Computer.

As Fig. 1 shows, the architecture of the hypertext is strongly interactive. It places the user in the position of a statistician, or at any rate of a user already trained in statistical methods. Learners are required to put their intuitions to full use. Users are introduced immediately to the particular problem that create for themselves and which they recognise and classifiy in the *BivarZero* initiation structure. The general sequence is strictly oriented towards the statistical solutions to the three possible problems that the user may encounter:

1. analysis of the relations between two quantitative variables;
2. analysis of the relations between two qualitative variables;

---

[3] *BivarZero* means that we are dealing with a prototype module (module 0) that tackles the measurement and analysis of bivariate dependence only for the purposes of demonstration. The experiment presented here concerns, in particular, the area of action relative to the dependence between two quantitative variables.

3. analysis of the relations between one qualitative and one quantitative variable.

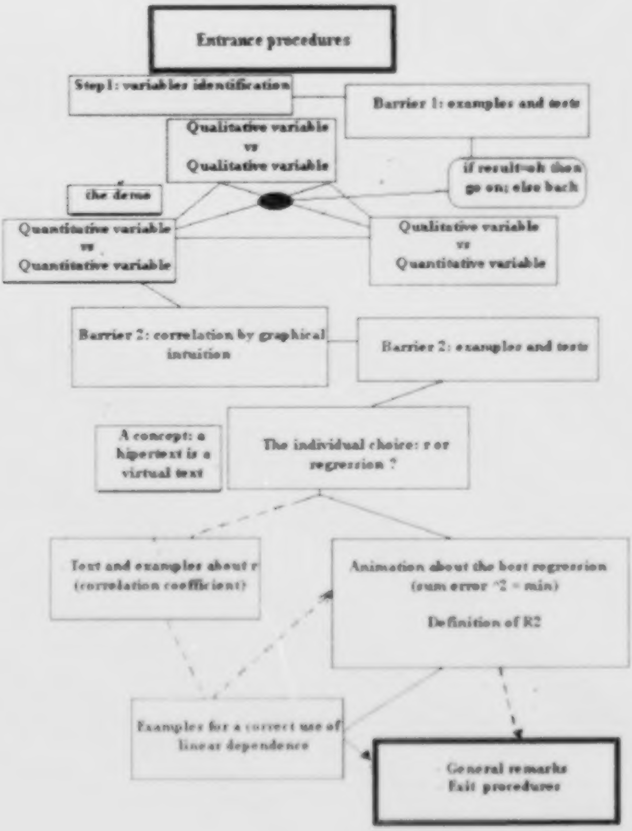

Entrance procedures
Step1: variables identification
Barrier 1: examples and tests
Qualitative variable vs Qualitative variable
if result=ok then go on; else back
the demo
Quantitative variable vs Quantitative variable
Qualitative variable vs Quantitative variable
Barrier 2: correlation by graphical intuition
Barrier 2: examples and tests
A concept: a hipertext is a virtual text
The individual choice: r or regression ?
Text and examples about r (correlation coefficient)
Animation about the best regression (sum error ^2 = min)
Definition of R2
Examples for a correct use of linear dependence
General remarks
Exit procedures

Figure 1. *The Hypertext Map and BivarZero journey*

In the introductory step, users are invited to pose their own concrete cognitive problem in a precise space of action. This involves the recognition of their problem in one of the three types of relation provided, by arranging the available information in the context of the measurement scales for the variables (Leti, 1983).

BEST COPY AVAILABLE

At this point, the path is directed by the classification tests that are applied to the concrete problem, but the connections among the three areas of action are many, and in any case the treatment of the basic subjects is common to all three. The subjects considered as basic are anchored to key words and comprise questions of univariate statistics, such as position and variability indices, and of mathematics, such as certain notions of analytic geometry and derivation.

The space of action actually presented in this demonstration is that relating to the dependence between two quantitative variables, and more specifically to linear dependency.

The hypertext experiment used in the demonstration was devised by guiding the user to the construction of a mental model for the correct use of the notion of linear dependence. The perturbations preliminary to this subject are essentially visual, and aim to arouse the graphic-geometric intuition of the subject (visual perturbation of the intuition).

The user, without any notions other than those provided in the path through the initial step, is requested to associate four forms of dependence to the same number of Cartesian graphs. Any error in this association is explained by a learning path relating to the reasons for this error, which involves navigation sideways and backwards along the basic concepts of univariate statistics and mathematics.

Once the barrier of the intuition perturbation step has been overcome, the user is able to proceed to the next level where two possible paths are introduced graphically and intuitively: 1) the synthetic measurement of the linear correlation between the two variables using the r coefficient of Bravais-Pearson; 2) the approximation of the relation between the two variables by means of a straight line, leading on to the estimation of the parameters of the regression line and the verification of the hypothesis of a linear model.

The two paths are obviously linked at numerous points and it is thus that users give life to the real text themselves: i.e. from a collection of information that is arranged only in a virtual manner into a text, the reader becomes the protagonist and actuator of the hypertext. The path actuated is clearly a function of the specific aim of the user and of the type of signification that he has given to the visual perturbations received in the previous step.

In a hypertext strongly directed to the solution of a specific problem such as the *BivarZero* experiment, it seemed very important to prepare a "help" image that allows the user at any moment to solve any problems of confusion or loss of way conected to the exploration of the intermediate levels or of the information base.

Our illustration of the potential of the *BivarZero* experiment loses much if it is presented in the form of a sequential text such as this paper. Consequently, the material here presented will be made available to anyone who contacts the authors.

4. Suggestions concerning the consultation strategies and the type of learning

BivarZero was demonstrated to a number of first year students of Business Economics at the University of Bologna in order to observe their behaviour. Though this was clearly not a true test of the functionality of the hypertext, a number of technical solutions were improved thanks to our observation of the student trial.

For example, a number of possible solutions were tested with reference to the need to provide perturbations to the intuition in the phase involving the first barrier. We noted that the different solutions we tried had a different impact according to whether the subject's background was technical/business high school or the more academic *liceo*. The fading effect applied to the explanatory picture or to the lower level chart was generally more successful among students from the *liceo*. The perturbation with interactive graphics or with a text to complete with the correct phrase worked considerably better for those subjects who, having a background of technical/commercial studies, interpreted the study of Statistics in a more practical manner.

The present demonstrational/experimental state of *BivarZero* has not enabled us to formulate any real tests of the reaction to the hypertext tool in the teaching of Statistics. Nevertheless, it is our intention to assess as soon as possible the two fundamental assumptions of *BivarZero*. Firstly, the non-causality of its consultation, which is intended to be dictated by the intuitive strategies of the user, and secondly, that the hypertext tool can be very useful in contexts such as a professional or company situation in which Statistics is considered as an operative tool for the solution of concrete problems.

5. Final considerations

The teaching of technical disciplines by means of recourse to multimedia tools is seen by many as the future for both scholastic and also professional teaching. In this context, two connected considerations seem to be the order of the day where there is authoritative debate on the subject (Gardner, 1993).

The first of these is that the hypertext and the other hypermedia educational tools will enable us to construct, in the near future, laboratory environments in which, through the use of recognition and programming software, every user/learner will increasingly be able to cut to his own shape a cognitive experience in which he is actively involved. In other words, the learning pattern of each user/learner will be increasingly a function of his specific requirements, in relation to a menu of subjects and events prepared by the authors of new learning products such as the hypertext.

The second consideration concerns the consequent transformation that the role of author/educator will undergo in the planning and supply of teaching. The increasingly virtual characteristics of the hypertext, and its quality of continuous encyclopaedic expandability, will lead authors to work increasingly in integrated multi-disciplinary groups, in which the role of author will lose some of its authority over the work itself. The task of the authors will thus be far more complex in that they will have to organise the learning in a collaborative manner, taking as much account as possible of the various potential paths/requirements that exist among the readers/users. On the other hand, only a group project will be able to ensure the modular nature of education and the interconnection between different phases and subjects (Adams *et al.*, 1989).

The hypertext and, more generally, hypermediality in teaching will then be able to reveal all its potential for the organisation and interdisciplinary fruition even of particularly technical and apparently unconnected subjects or disciplines such as those dealt with by Statistics.

Bibliography

Adams D., Carlson H. and Hamm M. (1989), *Cooperative learning*, Englewood Cliffs, New Jersey.

Antinucci F. (1991), Sulla natura dell'ipertesto, *Golem*, Anno III, n. 5.

Barrett E. ed. (1989), *The society of text: Hypertext, Hypermedia, and the social construction of information*, Mit Press, Cambridge MA.

Corcione D. and Di Tonto G. (1990), *Dal testo all'ipertesto: teoria, utilizzo e aree applicative*, Jackson, Milano.

Gardner H. (1993), *Educare al comprendere*, Feltrinelli, Milano.

Leti G. (1983), *Statistica descrittiva*, Il Mulino, Bologna.

Naddeo A. (1984), *Statistica di base*, Kappa, Roma.

Rondot F. (1993), I desideri del lettore multimediale, *Multimedia*, Anno I, n. 3.

Varisco B.M. (1991), Ancora sugli ipertesti, *Golem*, Anno III, n. 5.

STATREE: AN EXPERT SYSTEM FOR CHOOSING SUITABLE STATISTICAL DATA PROCESSING TECHNIQUES

Claudio Capiluppi and Luigi Fabbris
Dipartimento di Scienze Statistiche
Via S. Francesco 33, 35121 Padova, Italy

1. Introduction

The choice of a statistical method suitable for the analysis of a statistical dataset depends on several parameters: (i) the survey purposes, to be specified in terms of the research hypotheses and the population, (ii) the data formation environment, *i.e.* the survey and sample designs, the problems faced at the data collection and manipulation stages, (iii) the data measurement scale and the information on and/or the hypotheses about the distribution of variables.

Hierarchical decision trees showing the logical path which the informed researcher should go through have been drawn by many Authors (Andrews *et al.*, 1974; Harshbarger, 1971; Hays, 1973). The computer program STATREE is a personal computer guide for the selection of methods suitable for a directed analysis of data obtained in a given setting.

The system defines the statistical problem by iteratively acquiring from the user his/her aims and the information about the data at hand. The user is presented with a sequence of questions and a set of possible answers between which he/she is asked to choose. The system-user dialogue is assisted by an on-line statistical help accessible at any stage of the search.

The program can perform some analyses, provided the SAS statistical package is present on the user's computer. The user does not need to know how SAS works, nor its specific parameters. He need only to identify the dataset and the variables to process.

STATREE may be defined as an "expert system" because it fits many necessary attributes of a statistical expert system (Hand, 1985b), and in particular it allows the researcher to add new knowledge to the existing database.

2. The software program

The software program is composed of four components: the database containing the available statistical expertise, the structured query module to the database, the on-line help for statistical topics and technical terms, and the interface to the SAS data processing system.

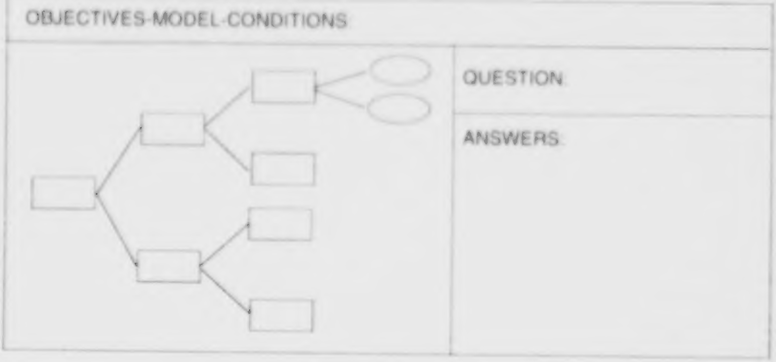
The knowledge of the system is based on a hierarchical network database, that is, a decision tree-like model (Fox, 1983). The database architecture shows some innovations with regards to previous implementations (Hand, 1985a) and allows for consistent updating of the meta-data base. Thus, the system can easily "learn" new techniques and new solutions to problems. The database creation and updating are enabled through an interactive fullscreen oriented managing module (Capiluppi, 1993).

STATREE is a program running on personal computers with Intel 80x86 processors and a DOS 5.0 operative system (Fabbris *et al.*, 1993). An 80286 personal computer with VGA graphic adapter and 1 MB of EMS is the minimal system configuration necessary to run STATREE. The program can also run in the Windows 3.1 environment. STATREE is written in C++ and SAS/SCL languages. The installation on hard disk needs about 1 MB of free space.

3. The query module

The interactive module for the selection of the suitable techniques shows a sequence of questions, in Italian, about the nature of the data and the scope of the analysis, and the user answers the questions by choosing one of the suggested items. The answer is selected using the arrow keys, <UP> and <DOWN>; the answer is then confirmed by hitting the <ENTER>, and the system goes to the next node of the decision tree. Hitting <ESC> the system goes back to the previous node.

The user interface of this module is made up of the following four windows.



The <F1> key shows the program help, that is the functions associated with the active function keys:

F1	program help
F2	context statistical help
F3	index of the statistical help
F4	references
F10	exit
ESC	back to previous node
↑ ↓	answer selection

4. The statistical help

The interactive selection process is provided with an on-line statistical help, supplying information about the technical terms and concepts in the questions. The help is organized into topics, from general to specific according to the tree structure, and implemented in help frames of variable length. Each question of the decision tree is associated with one such help frame by a keyword, isolating the argument pertinent to the context.

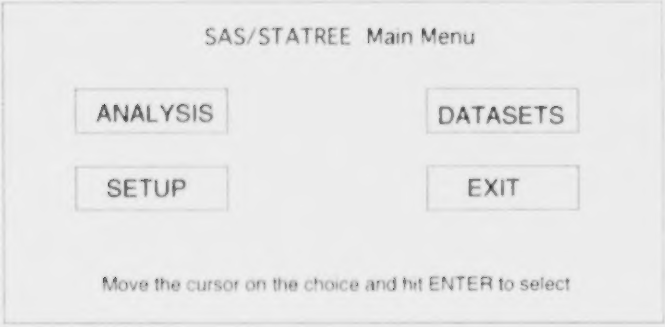
This context help is accessed by hitting the <F2> key, which opens a window on the screen displaying the help frame text. It is always possible to access any other topic through the general help index, by hitting the <F3> key, and then selecting the item of interest in the list.

The help text contains several references, which are available on-line in the system and accessible to the user through the <F4> key.

5. The interface to the SAS package

When a technique has been selected, the analysis can be performed by hitting the <ENTER> key. The analysis is carried out by the interfacing module to the SAS package. The user interface of this module is that of the SAS environment, and is based on dialogue windows and fields which allow the user an easy way to supply the necessary information to the system. Either the arrows or the function keys, or <TAB>, as well as a mouse, can be used to move through the windows fields, while <Fl> always opens the window help. The main menu displays the following four choices:

☐ ANALYSIS	carry out the selected analysis;
☐ DATASETS	build or modify the content of a dataset;
☐ SETUP	set up the SAS interface module;
☐ EXIT	return to the decision tree.


SAS/STATREE Main Menu

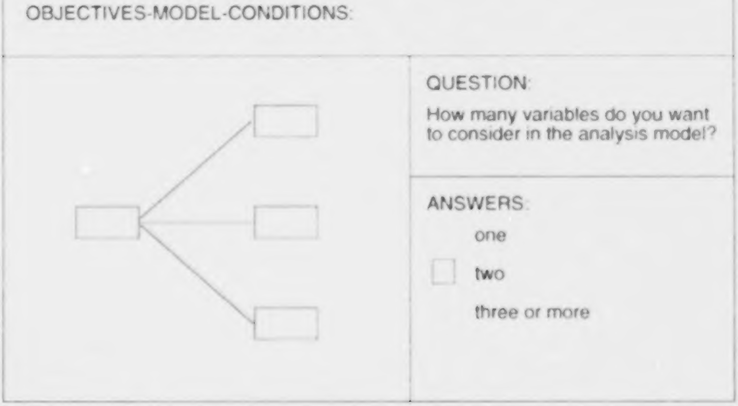
Move the cursor on the choice and hit ENTER to select

The module carries out the analysis by writing a suitable program (job) in the SAS language that is submitted to the SAS package. In this way, all computations are delegated to the SAS routines and procedures. The output is displayed in the OUTPUT window. The SAS program built by the system can be accessed by the user in the PREVIEW window.

6. A sample work session

Suppose we want to measure the relation between income and education of a set of subjects. The system acquires the problem description by asking the following questions:

Q1: How many variables do you want to consider in the analysis model? [A: two];



QUESTION:

How many variables do you want to consider in the analysis model?

ANSWERS:

one

☐ two

three or more

Q2: Which is the measurement scale of the variables? [A: one quantitative and one ordinal];

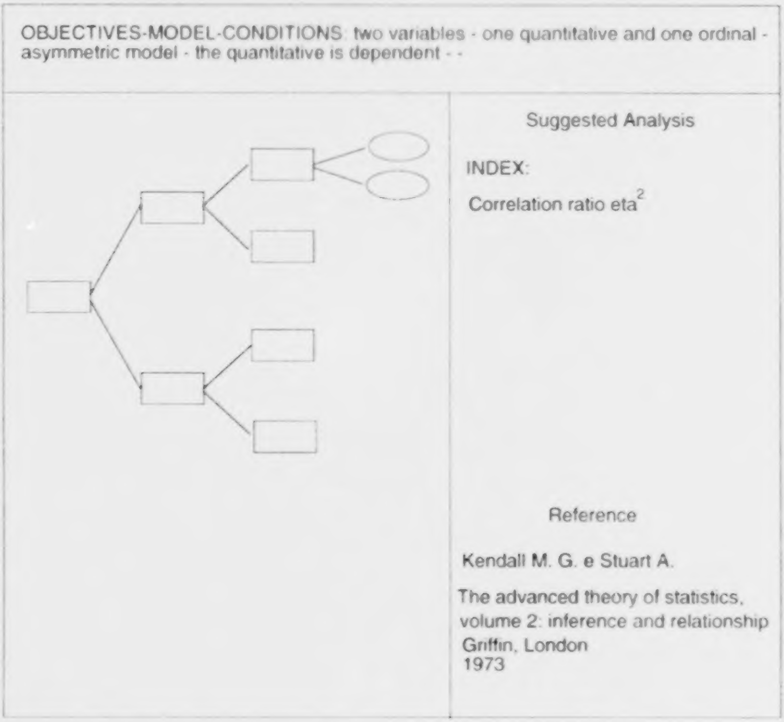
Q3: Is the model symmetric or asymmetric? [A: asymmetric];

Q4: Which one is the dependent variable? [A: the quantitative variable].

At this stage the system knows that the analysis is a bivariate asymmetric analysis with a quantitative dependent variable and an ordinal independent variable. Now it is possible to choose between two indices, the correlation ratio η^2 or the serial correlation ratio:

Q5: Do you want to consider the ordinal variable as a discrete class reduction of a normal variable? [A: no]

The suitable solution is then the correlation ratio η^2:



OBJECTIVES-MODEL-CONDITIONS: two variables - one quantitative and one ordinal - asymmetric model - the quantitative is dependent - -

Suggested Analysis

INDEX:

Correlation ratio eta^2

Reference

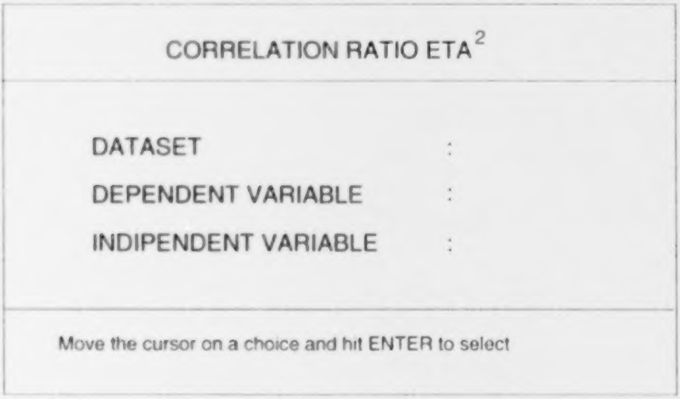
Kendall M. G. e Stuart A.

The advanced theory of statistics, volume 2: inference and relationship

Griffin, London

1973

Hitting <ENTER> activates the SAS interface module and displays its main menu. Choosing ANALYSIS the system gets the necessary inputs to process the data:



CORRELATION RATIO ETA^2

DATASET :

DEPENDENT VARIABLE :

INDIPENDENT VARIABLE :

Move the cursor on a choice and hit ENTER to select

Once the dataset and the variables have been specified, the system performs the analysis and displays the final output.

7. Perspectives

STATREE has been designed for teaching purposes. The iterative *dialogue* between user and system to identify the method suitable for the research problem at hand develops the user's ability to screen the properties of the methods available. So the system may be used as a self-teaching tool, with an occasional consultation of the manual

The program may be used also by researchers who need to use statistics but have a limited statistical background. It represents an operative tool for performing data analysis without experience of mathematical packages, and avoiding misuse of inappropriate analysis techniques.

Up until now, the statistical analyses supported in STATREE are the univariate and bivariate ones. Later versions will embed in the current tree the analysis of rates and proportions in longitudinal studies, and the main multivariate techniques.

Bibliography

Andrews F.M., Klem L., Davidson T.M., O'Malley P.M. and Rodgers W. (1974), *A guide for selecting statistical techniques for analyzing social science data,* ISR The University of Michigan.

Capiluppi C. (1993), *Implementation of the knowledge database of STA-TREE, an expert system for the selection of the statistical method for data analysis*, Statistic '93, Wollongong, NSW, Australia.

Fabbris L., Capiluppi C., Giancotti G. and Meneghello A. (1993), *STA-TREE 1.0 manuale per l'uso*, Summa, Padova.

Fox J. (1983), *The theory and practice of expert systems*, British Computer Society Specialist Group on Expert Systems, Newsletter No. 8 (May), pp. 14-16.

Hand D.J. (1985a), Choice of statistical technique, in *Proceedings of the 45th Session of the International Statistical Institute*, Amsterdam.

Hand D.J. (1985b), Statistical expert systems: necessary attributes, *Journal of Applied Statistics*, 12(1).

Harshbarger T.R. (1971), *Introductory statistics: a decision map*, The Macmillan Company, New York.

Hays W.L. (1973), *Statistics for the Social Sciences*, Second Edition, Holt, Rinehart and Winston, New York.

DESIGNING CURRICULAR SOFTWARE FOR CONCEPTUALIZING STATISTICS

Steve Cohen, George Smith, Richard Chechile and Robert Cook
Curricular Software Studio, Arena User Area
Tufts University Medford, MA 02155, USA

1. Introduction

Over the past decade, software has found its way into the curriculum for a variety of educational purposes. This has been especially true in the case of introductory statistics courses. There are at least two principal uses for technology in such introductory courses. The first is to free students of excessive computational responsibilities, thereby enabling them to spend more time analysing data and assessing results. The second is to help students to get a deeper understanding of the fundamental concepts taught in an introductory course, thereby putting the students in a better position to use powerful tool-based software and to interpret results. In a sense, both of these uses have the same aim, namely to improve the general quantitative reasoning skills of the students. In many respects this is the greatest challenge of an introductory statistics course, or of any course with a large quantitative component.

For years, mainframe tool-based software was used to serve both purposes. However, the promise of helping students to get a deeper understanding of fundamental concepts seemed specifically suited to the more interactive, easy-to-use programs that run on personal computers. A link between interactivity and deep understanding could be observed by watching how the small fraction of superior students in an introductory statistics class mastered concepts. Rather than simply memorising formulae, they would perform thought experiments, mentally clarifying and dissecting the concepts. They might look at a computational formula for a sample variance and wonder what would happen if the deviations were not squared. Posing these kinds of questions and discovering both answers and new questions seems to be the hallmark characteristic of students who leave an introductory statistics course with a facility for quantitative reasoning. Software that puts students in a position to pose questions and execute thought experiments on a computer might accordingly enable a much larger number of students to develop a deeper understanding of statistical concepts.

2. An initial design

In 1987, Tufts University's Curricular Software Studio, with support from FIPSE (Fund for the Improvement of Post-secondary Education), began to develop a set of microcomputer based programs, running under Microsoft Windows, that would permit students to explore concepts in introductory statistics. We decided that the programs needed to have the following design features and functions. Each program would cover a focused topic, such as Displaying Data, that typically takes one or two weeks in an introductory course. The programs would take advantage of a point-and-click graphical user interface that would allow students to access options and to execute commands easily. The pull-down menu options at the top of the program would contain a number of functions that would permit students to manipulate graphs, to examine data, and perform experiments. Students would have instant access to a variety of display options. The programs would offer discipline-specific data sets and random variables, so that students in the various discipline-specific statistics courses could encounter generic statistical concepts in a context with which they were familiar. After a year of development, we introduced seven programs into introductory statistics courses at Tufts, covering topics ranging from Displaying Data to Sampling. Each had a design similar to the Displaying Data program shown in Fig. 1.

We had very little success with this software. Only a small fraction of students, those that typically perform thought experiments and come away with a deep understanding without the aid of software, actually used the programs for an extended period of time. A certain number of the students would point and click, though without much purpose. Perhaps they expected that at some juncture a graph, table or result would appear on the screen that would help forge a useful picture of what had transpired. These students seemed willing to explore, but they were not capable of posing questions and executing experiments. Hence, they typically were not in a position to interpret a result. Finally, the majority of students did very little work with the programs. After a short time, they simply dismissed the opportunity to use the programs, mostly because it was not clear to them just what they were supposed to do.

We used the spring semester to examine our discouraging results and to reflect on the kinds of redesign that might prove useful. The software was originally written with the idea of giving students a simple-to-use platform for exploring a topic. The programs provided students with an environment rich in functionality, but with too little structure to permit those with sparse knowledge of the topic to pose questions or to explore successfully and intelligently. For instance, students using the Displaying

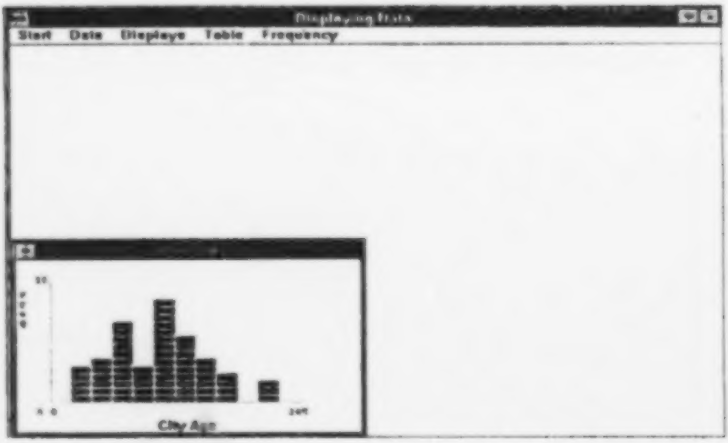


Figure 1

Data program shown in Fig. 1 might create the histogram shown in the bottom left hand corner. After they had created the histogram, we hoped they would perform the following four operations:

- Formulate a question about the histogram
- Determine which menu option best addresses the question
- Select the sub-option which best addresses the question
- Successfully interpret the result

For an excellent student, this is a challenging exercise. For a less than excellent student, it is often overwhelming. When the task becomes overwhelming, it appears that students often cease to explore, and active experimentation is replaced by a short session of button pressing and recording of results. All but the very best students seemed to fall back into a passive style of learning at the earliest opportunity. The challenge of redesigning the software was to provide students with enough structure to pursue an active, experimental style of learning without having the program tell them what to do. Direct instruction would allow students to assume a passive style of learning, and they would then be likely to consider the interactive experiments as a rote exercise. Our concern about this applied not only to the software, but to any curricular materials that went with it. Written assignments that dictated what buttons to press, what experiments to perform, and what results to record might also encourage a student to fall into the very passive style of learning that we were trying to eliminate.

3. Redesigning the software

We tried to address each of the four operations that seemed to stand in the way of students performing and successfully interpreting experiments. We saw two flaws in the original design that might make it difficult for students to pose a question. Firstly, they could perform any number of statistical operations without having a clear question before them to begin with. For instance, faced with a data table, students could create histograms, sort the data, consider relative frequency, etc. Narrowing down the number of choices seemed imperative. Secondly, even if the students did begin with a clear question, they might not see how to translate it into a program option.

To remedy these problems, we first decided to divide each program, and the functionality it offered, into a series of screens. Each screen would offer only a few options.

Some of the options would permit the students to perform an experiment about a specific concept, and others would permit the student to move to another screen and perform a new set of experiments. To help students focus on the particular concepts and set of experiments on a screen, we also provided a one or two sentence "scaffolding" at the top of the screen. For instance, a screen in the Displaying Data program might be devoted to examining how the number of intervals used to draw a histogram affect how data are lumped and consequently how they end up being displayed graphically.

In order to permit a smoother translation between an experimental option and a question, we presented all the options on a screen as questions. To follow the previous example, the program would prompt the student, "How many intervals would you like to try?", and then offer a set of options, posed as questions, including the opportunity to enter their own number.

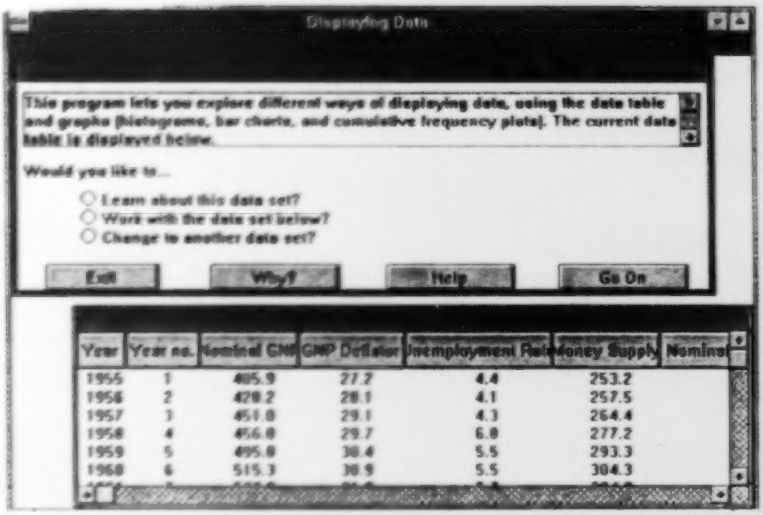
Finally, for each question, we offered both a WHY and a HELP button. When we originally introduced the software, we observed three fairly distinct patterns of response among those using it. Superior students tended to learn effectively from it from the outset, with no need of prompting. Others would use the programs, but without much purpose or sense of direction. Finally, most would simply give up shortly after they began. The WHY option is intended to help those who were having difficulty seeing a purpose. Hitting the WHY button when facing a choice produces a reason why the choice is an appropriate one to be facing. This usually takes the form of a one sentence statement of a typical consideration that someone might focus on when making the choice. For example, hitting WHY when hesitating over the question, "Do you want to change the number of intervals?", produces, "Maybe the histogram will take on a very different appearance with a different number of intervals."– just the sort of thought that a good teacher might whisper in the ear of a student who is hesitating in the middle of an experiment. We also hoped that the ubiquitous presence of the WHY button would induce students to begin asking themselves, rather than the program, why they might want to pursue each option. In the long run this might even yield a reflex action of posing questions and asking what experiments would be appropriate for answering them.

The HELP option is intended to rescue those who feel lost in trying to use the software to execute experiments. HELP responses are often one or two paragraphs aimed at helping the student to understand the goal of the experiment. Sometimes the text includes instructions, detailing how to carry out specific experiments yielding clearly interpretable results. But even in these cases the student is actively eliciting the information, looking for specific things that will help them to take the next step, in much the way that a superior student will take a quick look at one or two pages of a book when working something out in a thought experiment.

4. ConStatS: the redesigned package

The first screen of the redesigned Displaying Data program is shown in Fig. 2.

The scaffolding at the top orients the student. There are three options, posed as questions. For any of the options, the student can ask WHY, or select HELP. The first option, learning about the data set, is a simple experiment. Students selecting this option will see a short description about the data, why it was collected, etc. The other options take students to different screens. These choices lead to different pathways through the program, pathways that often converge into one another. These pathways



This program lets you explore different ways of displaying data, using the data table and graphs (histograms, bar charts, and cumulative frequency plots). The current data table is displayed below.

Would you like to...

○ Learn about this data set?
○ Work with the data set below?
○ Change to another data set?

Year	Year no.	ominal GN	GNP Deflator	nemployment Rat	oney Suppl	Nominal
1955	1	405.9	27.2	4.4	253.2	
1956	2	420.2	28.1	4.1	257.5	
1957	3	451.0	29.1	4.3	264.4	
1958	4	456.8	29.7	6.8	277.2	
1959	5	495.8	30.4	5.5	293.3	
1960	6	515.3	30.9	5.5	304.3	

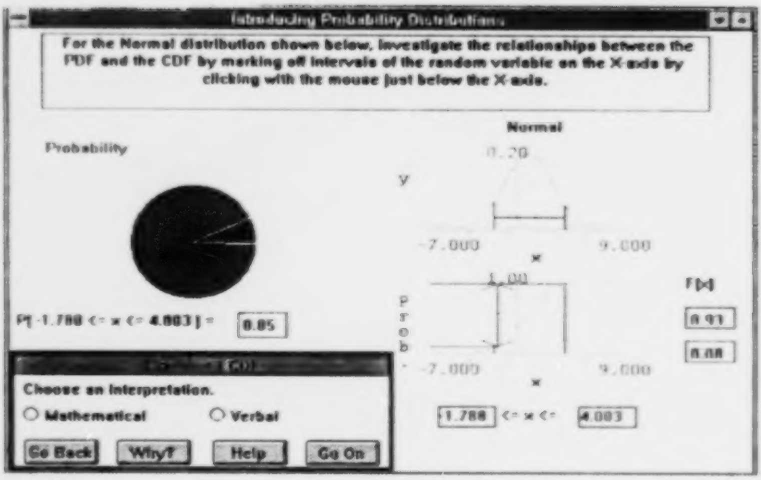
Figure 2

provide an unobtrusive built-in structure that helps to guide the student along in an orderly fashion. More guidance is provided at some places by making the decision one between a default value offered by the program and a value of the student's own choosing. Each screen has a "Go Back" option so that the student can always back up along a pathway to review or reconsider earlier choices. The students are always in control, not just in the sense that they choose what to do next, but in the sense that nothing ever happens on the screen except through choices that they make.

Fig. 3 shows an experiment screen from the Probability Distributions program. It enables students to examine how probability distributions convey probability, and to compare the PDF and the CDF. Clicking on the abscissa of either the PDF or CDF will update the highlighted interval on both distributions. The corresponding probability is illustrated in the pie chart to the left. Students do not arrive at this screen until they have selected a distribution for study, and have examined how varying parameter values affects the form that the distribution takes on.

In all, ConStatS consists of nine independent programs that offer about 15 to 20 hours of curricular time for the first 7 to 10 weeks of an introductory course. As redesigned, ConStatS embodies the view that microcomputer-based curricular software is truly effective when it functions as a medium in its own right, entirely different from classroom

BEST COPY AVAILABLE



Introducing Probability Distributions

For the Normal distribution shown below, investigate the relationships between the PDF and the CDF by marking off intervals of the random variable on the X-axis by clicking with the mouse just below the X-axis.

P[-1.788 <= x <= 4.003] = 0.85

Choose an interpretation.

○ Mathematical ○ Verbal

-1.788 <= x <= 4.003

Figure 3

lecture, drill, textbook reading, and educational television. None of these others can be devoted exclusively and uncompromisingly to active experimentation. Carefully designed curricular software can be.

5. Using and evaluating ConStatS

ConStatS is used in almost all the discipline-specific introductory statistics courses taught at Tufts. Even though most students seem able to use the software for an extended period of time without a great deal of direction, most of the instructors provide assignments that provoke open-ended use of the software. These assignments usually contain short essay questions that address the kinds of conceptual issues captured in ConStatS experiments.

During the past year, ConStatS has been at the center of a large scale, FIPSE funded project for evaluating the effectiveness of curricular software. The goal of the evaluation was to assess if ConStatS in particular, and curricular software in general, could help students to develop a deeper conceptual understanding of statistics (or of any other subject matter built around a comparably abstract theory).

The evaluation took place in the fall 1992 and the spring 1993 semesters. Classes at three universities participated. In total, seven classes with 303 students participated as experimental groups, and two classes (from Tufts) with 49 students participated as control groups. Over 200 experimental subjects came from the introductory statistics course for psychology students at Tufts. Other subjects were enrolled in statistics courses for biology, education, child study and economics. The main control group was taught by a member of the team that designed ConStatS. Both control groups used tool-based statistics software for mainframe and personal computers, but did not use ConStatS.

The principal assessment instrument was a 103 item test of conceptual understanding that was derived from ConStatS. Questions were scored on a five point scale: 1-no understanding, 2-little understanding, 3-moderate understanding, 4-good understanding, 5-excellent understanding. We computed average results for each question, and then computed difference scores by subtracting the results of the control group from those of the experimental groups:

- The overall experimental group did better on 82 out of 103 questions
- The Tufts experimental group did better on 89 out of 103 questions

The 10 questions on which the experimental group showed the greatest differential involved transformations (z scores and linear transformations), probability, the concept of deviation, and the sensitivity of summary measures. In many respects these results need to be conditioned on exactly what parts of ConStatS were used in each class, and on the kinds of experiments students performed. The version of ConStatS being used in the evaluation program has a facility that records just how each student interacts with the software, and the specific experiments they do. An analysis of all this trace data from the evaluation is in its infancy and beyond the scope of this paper. However, the comprehension test results, and our ability to connect them with specific screens, have already helped us to add or to redesign 20 screens, and to make a number of small but significant changes. We will repeat the evaluation next year, and learn then if we have improved the software in the way we have intended.

Two other preliminary results are worth noting. On many questions, the patterns of the errors were very revealing. For instance, when students were asked "What is it that makes a variable a random variable?" most students gave one of three wrong answers, indicating specific misunderstandings in the absence of correct comprehension. These results will be of value independently of the software. Secondly, when we scored the experimental group results using three categories (2 or less, 3, 4 or greater), 74 out of 103 questions had bimodal distributions. Students appeared to get either a fairly strong or a fairly weak understanding of most concepts. The trace data, we expect, will indicate which students fall where, and why.

THE USE OF A COMPUTER BASED ADVISORY SYSTEM AS A TEACHING AID IN STATISTICS

Joe Griffin and Mary Rouncefield
Chester College of Higher Education
Cheyney Road, Chester, CH1 4BJ, UK

1. Knowledge based systems and learning

There has been much research into the use of artificial intelligence in computer aided learning. One area that has already shown promise is the use of knowledge based systems (KBSs). These systems, which attempt to emulate the decision making processes of an expert in a particular subject domain, have a number of potential applications in education.

Briggs (1988) has suggested five ways in which they might be used:

1. *As curriculum resources*

The ability to use knowledge based systems to store the expertise of humans should imply that it could be possible to store, for example, the knowledge of an expert in physics or medicine and to make this available to all students. However, not all knowledge or expertise can be expressed using a rule based methodology but it may be possible to use KBSs to store some parts of our knowledge.

2. *As part of intelligent tutoring systems*

Intelligent tutoring systems (ITSs) are computer programs that attempt to emulate the learning situation by representing knowledge about the subject matter, teaching styles and learner knowledge. The way in which knowledge is represented is of fundamental concern, and KBSs offer one solution. However, although fairly simple in concept, there have been, and still are, many problems with ITSs, especially in the construction of adequate models of the ideal learner and in the assessment of the understanding of the user of the system.

3. *KBSs as simulators*

Simulations have existed in a variety of different guises, from role play to games, but computer simulations tend to share the common feature that they have an inability to explain their reasoning. With KBSs, the

BEST COPY AVAILABLE

learner can see how the conclusions have been arrived at, and the systems can explain all the stages between input and output.

4. *Principles of knowledge based systems*

KBSs are playing an increasing role in industry and commerce and are now considered one of the 'mainstream' software applications. As such, it is now a truism that many of those leaving our colleges will come into direct contact with KBSs. Teaching the principles of knowledge based systems in the education sector is now a valid subject in its own right.

5. *Exploratory tools*

Finally, KBS shells can be used as tools to explore other subjects by allowing the students to act as experts (or to become experts if they don't know enough about a domain) and to restructure their knowledge using these systems. (Galpin, 1989; Griffin, 1992; Lippert, 1989; Webb, 1990). The process of knowledge engineering by which subject knowledge is restated in a suitable form for the KBS could be a powerful learning process

2. The possible need for a statistical advisory system

All students at Chester college take two main subjects in their degree courses. At present most undertake a large project or dissertation in both these subjects in their final year. For many of them this includes either some field work or experimental work that generates data for analysis and interpretation. Subjects such as Psychology, Biology, Geography, Sports Studies and Health and Community Studies all generate small research projects of this kind.

Many of these Science and Social Science students have had no formal teaching in statistics and most have not studied mathematics beyond the age of 16.

In order to help those students with their problems, a statistics lecturer has set up a STATISTICS CLINIC. This is well attended and at peak times there are more students needing help than can be sensibly dealt with. The statistics lecturer therefore decided to investigate the possibility of providing help for these students via a computer based advisory system.

The main question asked by students attending the statistics clinic is: WHAT DO I DO WITH THESE RESULTS? (How to do it is sometimes their second question). This has determined the scope of the proposed system.

The potential users of the system include all undergraduate students.

Different students will have varying amounts of computing expertise. In order to ensure that the proposed system was suitable for this group certain requirements were established. These requirements are:

1) It must be easily accessible by students with limited experience of computers, and
2) It must use simple language. Any technical terms must be explained to users as they go along.

Both of these requirements influenced the design of the user interface. It was assumed that students:

(a) either already know how to do the calculations involved (or can look this up in text books), or
(b) have access to a computer package that will do the necessary calculations.

As a result of these design considerations it was decided to investigate the use of knowledge based systems as a tool in which to develop the advisory system.

3. Development of the advisory system

A final year student (Carpenter, 1992) taking mathematics and computer studies at Chester College undertook the development of this advisory system, choosing statistical tests as her final degree project. The client who specified the requirements of the system was the lecturer in statistics at the college. The student used the Leonardo KBS development environment.

The development process itself proved to be an invaluable learning experience for that student. Together she and the lecturer had to develop the logical structure of the system, choose the order of the questions to be asked and plan the pathways users could take through it. The lecturer acted as the expert and the student as the knowledge engineer.

The resulting knowledge based system initially asks the user to think about what they want to find out from the data, i.e. what is the question being asked? Help is given in deciding what kind of data is available, and the level of measurement used, i.e. nominal, ordinal or interval. The KBS then advises the student on which simple statistical technique might be used. The system is designed with our particular clientele in mind and suggests only simple analyses.

4. Using the advisory system

Once the KBS had been fully developed and tested for reliability, it was decided to evaluate its use as an aid to solving problems. The target group of students were undergraduates studying subjects such as Physical Education, Psychology, Biology and Geography. The majority of these students have only studied mathematics to age 16 and none of them had studied it as part of their degree course. These students take an optional course in introductory statistics that makes considerable use of computer software to teach simple EDA techniques, non-parametric and parametric tests.

Evaluating the system

Students taking the first year "Statistics for Scientists" optional course were split into two groups. Both groups were given a series of three problems to consider with reference to a data set provided. They were not asked to do any calculations but to correctly identify which statistical analysis was required to help answer each problem. This was both a learning task and an assessment task. Students were told that the exercise was designed to help them in their revision (but their performance would not affect their grades).

Group A consisted of students who had to complete the task using their own lecture notes, and text books provided for this purpose. Group B consisted of students who had access to the KBS at individual work stations.

Both groups of students performed equally well on the three problems. Their mean scores on a total of 9 questions (3 per problem) are shown in Table 1.

	Mean score out of 9	Mean number of techniques correct
Group A	5.07	1.78
Group B	4.60	1.35

Table 1

There were no significant differences between these mean scores. As Group A students were able to refer to their own notes as and when they wanted to, and Group B students were faced with a computer system they had never used before, it is encouraging that the latter group were able to gain a similar degree of support from the advisory system.

5. Future developments

The pilot project just outlined, raised a number of interesting questions that have suggested further areas of improvement.

The development of a more user friendly interface (possibly WINDOWS based) to which the users may more easily relate, is one such area. Comments from the subjects suggested that this would be advantageous.

New users of the system will also need a tutorial on how to use the knowledge base with a full explanation of the facilities offered. Once these amendments have been made, it is proposed that the evaluation will be repeated again. We would expect that there will be a further improvement in the results achieved by users of the knowledge base and we eventually plan to install this on the college computer system thus making it available to all students.

Bibliography

Briggs J.H. (1988), *Learning With Expert Systems*, the Advisory Unit for Microtechnology in Education, Hatfield, U.K.

Carpenter N. (1922), *A Knowledge Based Statistical Advisory System*, Unpublished Dissertation, Chester College, England.

Galpin B. (1989), *Expert Systems in Primary Schools*, British Library Research Paper 73, British Library, London.

Griffin J. (1992), *Using Expert Systems to Enhance Learning*, Unpublished MSc Thesis, University of Keele, England.

Lippert R. (1989), Expert Systems, Tutors, Tools and Tutees, *Journal of Computer Based Instruction*, 16(1), pp. 11-19.

Webb M. (1990), Learning by Building Computer Based Qualitative Models, *Computer Education*, 66, November, pp. 27-32.

A COMPUTER AIDED SOFTWARE FOR TEACHING STATISTICS

Michel Janvier
ERES, Université Montpellier II, Place Eugène Bataillon
34000 Montpellier, France
Gilbert Saporta
CNAM, 292 rue Saint Martin, 75003 Paris, France
Eric Verdoire
CNAM, 11 Avenue du Docteur Grasset, 34000 Montpellier, France

STATCNAM is a software product designed for the teaching of statistics at the first level of technical studies at CNAM: The Conservatoire National des Arts et Métiers is the largest institution in France for continuous education with 100 000 students and a network of 50 regional centres. Several thousands of students are registered each year for the credit "statistical techniques".

The software may be used either

- for distance learning
- for self education
- as a complement of traditional teaching

An intensive use of graphics and simulation is done, and the software includes an analysis of answers.

It is divided into 6 chapters:

1. exploratory data analysis
2. simple linear regression
3. discrete probabilities
4. random variables
5. sampling
6. statistical tests

Chapter 2 is available at publisher STUD-I, Chapters 3 and 4 are in development.

The production of the software is made in association with Université Montpellier II, Unité de Biométrie INRA ENSAM USTL of Montpellier, and the Centre Régional Languedoc Roussillon of the CNAM by a team of teachers and scientists.

A task force defines the educational goals to fulfill and validates the product during each phase of its realisation.

This communication will present the various features of the software which runs on PC compatible computers under DOS or WINDOWS.

The course is not limited to a traditional presentation, it includes many graphical representations. A graphing assistant is part of the software as well as an integrated pocket computer.

A ROLE FOR COMPUTER INTENSIVE METHODS IN INTRODUCING STATISTICAL INFERENCE

Peter Jones, Kay Lipson and Brian Phillips
Swinburne University of Technology
John Street, Hawthorn 3122, Australia

1. Introduction

Introductory statistics courses have become increasingly prevalent in a wide variety of undergraduate and graduate programs over the past few years. This has resulted in the study of inferential statistics becoming the norm rather than the exception. Whilst statistics was once a course chosen by the more mathematically able student, many current students of statistics have little mathematical aptitude or expertise (Tanis, 1992 for example). As a result, many introductory statistics courses have moved away from including much of the statistical theory that underpins inference to become basically technique oriented "recipe book" courses. The inherent danger of producing students capable of performing various complex statistical tests without really knowing what they are doing is obvious, particularly if the course includes the use of a sophisticated statistical computer package.

To overcome the problem of having to teach inferential statistics to mathematically unsophisticated students it has been suggested by several statistics educators that the appropriate use of the computer in the classroom may help. In particular, computer based simulation is seen as a valuable tool, for instance:

> Simulation can help convey both the hard idea that random variation has a pattern in the long run and specific facts such as the central limit theorem (Moore, 1992, p. 23).

and

> Teachers can use simulation to illustrate ideas that are not otherwise accessible to beginning students. For example, there are many proofs of the central limit theorem, and many of them are short. But none are particularly intuitive or accessible to students who know nothing of moment generating functions. On the other hand, it is easy to demonstrate the central limit theorem, making the underlying definitions clearer and illustrating methods of simulation in the process (Thisted and Velleman, 1992, p. 49).

There is, however, a lack of substantive research into the effectiveness of simulation exercises and opinion is divided amongst statistics educators as to the advantages of such activities. As noted by Hawkins:

> ICOTS 2 delegates were treated to "101 ways of prettying up the Central Limit Theorem on screen", but if the students are not helped to see the purpose of the CLT, and if the software does not take them beyond what is still, for them, an abstract representation, then the software fails (Hawkins, 1990, p. 28).

Although untested, it would seem that exposing students to large numbers of computer simulations does not necessarily clarify for them some of the more fundamental issues in statistics. It may well be that this apparent lack of effectiveness is due to our lack of knowledge of how to teach effectively with computer technology and this is not surprising, given our relative lack of experience and theory to guide us. One area in which some theoretical work has been done is in determining the features of the software that may in theory prove to be more effective, such as the use of multiple representations, and dynamic interactive displays (Hawkins, 1990; Kaput, 1992; Rubin *et al.*, 1988) It is not the purpose of this paper to delve into the reasons why the particular software being used may or may not be effective, although this is in itself an extremely important issue. Our purpose is to suggest a rationale as to why computer based simulations are not as helpful as we might suppose, and to propose an alternative path leading to statistical inference, which potentially avoids this problem.

2. Statistical inference and the sampling distribution

A critical step in developing the theory of statistical inference is the idea of a sampling distribution – the recognition that the estimates of a population parameter will vary and that this variation will conform to a predictable pattern. Yet, for all its importance, experience and research have shown that the idea is generally poorly understood (Moore, 1992; Rubin, 1990, for example). One reason for this might be the way in which the idea has been traditionally introduced in statistics courses, using a deductive approach based on probability theory (Johnson and Bhattacharyya, 1987; Mendenhall *et al.*, 1990 for example). Such explanations are usually expressed in highly mathematical language, which tends to make the argument largely inaccessible to all but the more mathematically able. But perhaps more importantly, it is a theoretical development that is difficult to relate to the physical process of drawing a sample from a population. Statistics educators have come to recognise that there are deficiencies with a purely theory based explanation, and now often accompany or replace this with an empirical argument. The alternative interpretation uses the long run relative frequency approach to probability, where the sampling distribution is viewed as the result of taking repeated samples of a fixed size from a population and calculating the value of the sample statistic for each (Devore and Peck, 1986; Ott and Mendenhall, 1990, for example). The empirical approach has the advantages of being more readily related to the actual physical process of sampling and of requiring minimal use of formal mathematical language.

Because the computer has an obvious role in this empirical development of the idea of a sampling distribution, by enabling the repeated sampling to be carried out and summarising the results, a number of instructional sequences have been developed built around these capabilities (Norusis, 1988, for example). These approaches have been widely promoted and are now commonplace activities in introductory statistics courses. In our experience, however, they are less successful than statistics educators might have hoped. One reason for this may be the problem that students have when endeavouring to integrate their empirical experience of the sampling distribution with the theoretical model of the sampling distribution that is used in classical statistical inference. Eventually, in the classical approach, students must accept that the sampling distribution of the statistic of interest may be modelled by a known probability distribution, and that knowledge of this probability distribution is the basis for further estimation and hypothesis testing.

3. Using computer intensive methods

An alternative path to inference is provided by the developing area of Computer Intensive Methods. Basically, these methods are called "computer intensive" because they involve the computation of the statistic of interest for many data sets. According to Diaconis and Efron:

> The payoff for such intensive computation is freedom from two limiting statistical factors that have dominated statistical theory since its beginnings: the assumption that the data conform to a bell-shaped curve and the need to focus on statistical measures whose theoretical properties can be analysed mathematically (Diaconis and Efron, 1983, p. 96).

For the practising statistician, there are often definite advantages of using a computer intensive method rather than a classical method when restrictive assumptions about the nature of the data do not hold, or traditionally used test statistics are not appropriate. From the pedagogic point of view, there may also be advantages to the student of statistics.

Essentially, applying a computer intensive method means that the entire inference problem is dealt with from an empirical perspective. Students are not required to make the conceptual connection between the empirical and theoretical sampling distributions, but actually use an empirical sampling distribution as the basis for the inference. Consider the following example, taken from a typical introductory statistics text by Weiss and Hassett:

> A highway official wants to compare two brands of paint used for striping roads. Ten stripes of each paint are run across the highway. The number of months that each stripe lasts is given below.
>
> | **Brand A** | | **Brand B** | |
> |---|---|---|---|
> | 35.6 | 36.1 | 37.2 | 36.4 |
> | 37.0 | 35.8 | 39.2 | 37.5 |
> | 34.9 | 34.9 | 37.2 | 40.5 |
> | 36.0 | 38.8 | 38.8 | 38.2 |
> | 36.6 | 36.5 | 37.7 | 36.6 |
>
> Based on the sample data, does there appear to be a *difference* in mean lasting time between the two paints? Use a = 0.05 (Weiss and Hassett, 1987, p. 427).

Classical method of solution

From the question, the students must recognise that they are required to perform a hypothesis test for the equality of means with a two-sided alternative thus:

$$H_0: \ \mu_1 = \mu_2$$
$$H_1: \ \mu_1 \neq \mu_2$$

If we assume that the two samples are taken independently from two normally distributed populations with means μ_1 and μ_2 respectively, and the respective standard deviations are unknown, then a hypothesis test for comparing the means can be carried out by consideration of the test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}} \ .$$

which has a distribution which can be approximated by a *t*-distribution with degrees of freedom given by

$$df = \frac{(s_1^2 / n_1 + s_2^2 / n_2)^2}{\frac{(s_1^2 / n_1)^2}{n_1 - 1} + \frac{(s_2^2 / n_2)^2}{n_2 - 1}}.$$

where n_1 and n_2 are the sizes of the samples, $\bar{x}_1$ and $\bar{x}_2$ are the means of the samples, and s_1^2 and s_2^2 are the sample variances. Substituting in these formulae gives a value of the test statistic that is then compared with a value from the tables[1].

Rather than using a calculator to perform these calculations, such procedures are generally carried out using a standard statistical package such as Minitab. This is most likely how the student would be expected to proceed:

```
MTB > TwoSample 95.0 'BrandA' 'BrandB';
SUBC>  Alternative 0.

TWOSAMPLE T FOR BrandA VS BrandB
        N   MEAN   STDEV  SE MEAN

BrandA 10   36.22   1.14   0.36
BrandB 10   37.98   1.33   0.42
95 PCT CI FOR MU BrandA - MU BrandB: (-2.93, -0.59)

TTEST MU BrandA = MU BrandB (VS NE): T = -3.18 P = 0.0055 DF = 17
```

Of course the student must now relate the computer output to the original question, which would be, on the basis of the *P*-value, to reject the null hypothesis and to conclude that there is a difference in mean lasting time between the two brands of paint.

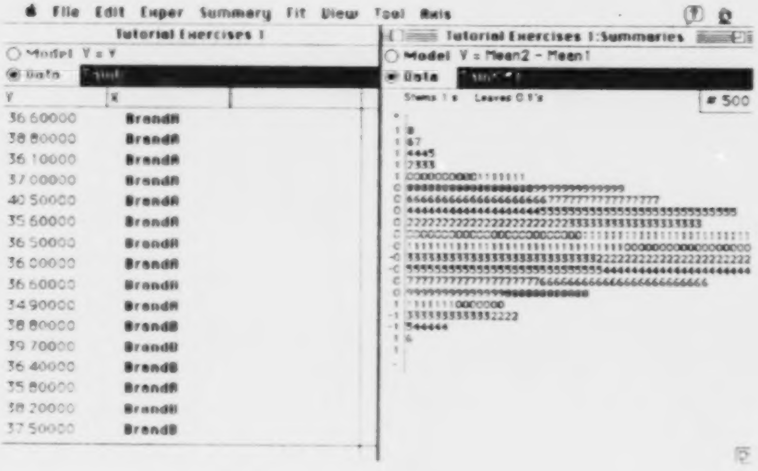
Computer intensive method of solution

The observed mean difference in the time that the paint lasts for the two brands is 1.76 months. This difference may be due to a real difference in the lasting time of the paints, or it might well have occurred by chance. Consider a null hypothesis that there is no relationship between the brand of paint and the number of months it lasts, so that all twenty observations essentially come from the same population. If this is true, then the observed difference in the means would not be particularly unusual when compared to many values of this mean difference that could be found by randomly selecting two groups of these sizes from the population. The computer intensive procedure used here is to divide the twenty observations into two equal sized groups at random, designating one as Brand A and the other as Brand B, and computing the difference in sample means. This is repeated many times, and as a result an empirical distribution for the difference in means is generated. The probability that there is a real difference in the mean may then be estimated by determining the proportion of the randomisation samples that give a difference as far from zero as the actual difference observed. Further details regarding this procedure can be found in Noreen (1989).

---

[1] The formulae given are those that are used when we cannot assume the variances of the two populations are equal. The assumption of equal variances was often made in the past because the distribution of the test statistic had otherwise only an approximate (*t*-distribution). The common use of computers has meant that this is no longer an issue. Whether or not the assumption of equal variances is made the classical methods of solution are very similar.

There are many computer packages readily available that enable the problem to be solved in the manner described. The empirical distribution of the difference in group means for the example under consideration shown in Fig. 1 was generated by *Models 'n' Data* (Stirling, 1991), a Macintosh based computer package designed for teaching and learning statistics which offers the facility of using various computer intensive methods:


File Edit Exper Summary Fit View Tool Axis
Tutorial Exercises 1
○ Model Y = Y
● Data
Tutorial Exercises 1:Summaries
○ Model Y = Mean2 − Mean1
● Data
Stems 1's Leaves 0.1's
# 500

Y	X
36.60000	BrandA
38.80000	BrandA
36.10000	BrandA
37.00000	BrandA
40.50000	BrandA
35.60000	BrandA
36.50000	BrandA
36.00000	BrandA
36.60000	BrandA
34.90000	BrandA
38.80000	BrandB
39.70000	BrandB
36.40000	BrandB
35.80000	BrandB
38.20000	BrandB
37.50000	BrandB

Figure 1

In this example, an empirical distribution of 500 observations was generated, and displayed as a stem and leaf plot. From the plot it may be seen that there was only one value of the mean difference as far or further from zero than 1.76, giving our estimated *P*-value as 1 in 500 or 0.002. Thus we see that such a difference in means is very unlikely to have been caused by chance alone, and conclude that there is a real difference in the mean lasting times of the two brands of paint.

Comparison of classical inference and computer intensive methods

The essential components of a hypothesis test have been well described by Noreen:

> Three ingredients are usually required for a hypothesis test: a hypothesis, a test statistic, and some means of generating the probability distribution of the test statistic under the assumption that the hypothesis is true (Noreen, 1989, p. 2).

It is in this process of defining a test statistic and determining its distribution that the essential difference between the classical and computer intensive methods lies. The test statistic and its distribution are often complex, and their origins obscure to students of classical statistical methodologies. The essential purpose of the process, to determine the likelihood of such a set of outcomes under certain assumptions, becomes easily lost. In classical hypothesis testing, the fundamental logic of the hypothesis test becomes entangled with the theory underlying the sampling distribution of the test statistic. Using a computer intensive method, an intuitively obvious test statistic is selected for which an empirical distribution is generated. The students remain linked to the overriding purpose of the exercise – the inference. Using the technology in this way amounts to forming a partnership between the student and the computer in which the computer takes on the lower level tasks of performing the numerous of calculations, whilst the student undertakes the higher order tasks of actually applying the logic of inference to the particular situation. In fact, the student must keep in mind the overall purpose of the exercise in order to carry it out successfully. Whilst the student of classical inference needs to recognise certain situations, and to make the appropriate assumptions, the students of computer intensive methods have been equipped with valuable problem solving skills that may have application in circumstances far removed from the one at hand.

The question arises as to the long term effect of the technology on the students' understanding of statistical inference when computer intensive methods are used. It is reasonable to contend that the act of carrying out hypothesis testing using computer intensive methods may actually enhance students' understanding of inference, a long term effect of the technology. It may even be that this deeper understanding could then be applied to classical hypothesis testing when students are exposed to this procedure, although one might ask why, in this computer age, this would be necessary.

Thus, it seems possible that using computer intensive method may enable beginning students of statistics, especially those without strong mathematical backgrounds, to appreciate the important principles of statistical inference. Whilst at this stage little research exists to support this theory, what does exist seems encouraging (Simon *et al.*, 1976, for example). A challenge for statistical educators in the 1990's is to determine if this is indeed the direction we should be following.

Bibliography

Devore J. and Peck R. (1986), *Statistics*, West Publishing Company, St. Paul.

Diaconis P. and Efron B. (1983), Computer-Intensive Methods in Statistics, *Scientific American* (May), pp. 96-108.

Hawkins A. (1990), Success and Failure in Statistical Education - A UK Perspective, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 1, pp. 24-32, .

Johnson R. and Bhattacharyya G. (1987), *Statistics Principles and Methods*, John Wiley & Sons, New York.

Kaput, J.J. (1992), Technology and Mathematics Education, in D. A. Grouws (eds.), *Handbook of Research on Mathematics Teaching and Learning*, MacMillan, pp. 515-555.

Mendenhall W., Wackerly D.D. and Scheaffer R.L. (1990), *Mathematical Statistics with Applications*, 4*th* edition, PWSP–Kent Publishing Company, Boston.

Moore D.S. (1992), Teaching Statistics as a Respectable Subject, in F. Gordon and S. Gordon (eds.), *Statistics for the Twenty-First Century*, MAA Notes No. 26, Mathematical Association of America, pp. 14-25.

Noreen E. W. (1989), *Computer Intensive Methods for Testing Hypotheses: An Introduction*, John Wiley & Sons, New York.

Norusis M.J. (1988), *SPSS/PC+ Studentware*, SPSS Inc., Chicago.

Ott L. and Mendenhall W. (1990), *Understanding Statistics*, 5*th* edition, PWSP-KENT Publishing Company, Boston.

Rubin A., Bruce B., Roseberry A. and Du Mouchel W. (1988), Getting an Early Start: Using Interactive Graphics to Teach Statistical Concepts in High School, in *Proceedings of the Section on Statistical Education*, American Statistical Association, pp. 6-15.

Rubin A., Bruce B. and Tenney Y. (1990), Learning about sampling: Trouble at the core of statistics, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 2, pp. 314-319.

Simon J.L., Atkinson D.T. and Shevokas C. (1976), Probability and Statistics: Experimental Results of a Radically Different Teaching Method, *Mathematical Education*, November, pp. 733-739.

Stirling D. (1991), *Models 'n' Data*, Intellimation, Santa Barbara.

Tanis E.A. (1992), Computer Simulations to Motivate Understanding, in F. Gordon and S. Gordon (eds.), *Statistics for the Twenty-First Century*, MAA Notes No. 26, Mathematical Association of America, pp. 217-225.

Thisted R.A. and Velleman P.F. (1992), Computers and Modern Statistics, in D.C. Hoaglin and D.S. Moore (eds.), *Perspectives on Contemporary Statistics*, Mathematical Association of America, pp. 41-53.

Weiss N. and Hassett M. (1987), *Introductory Statistics*, *2nd* edition, Addison-Wesley, Reading, Massachusetts.

CHEMOMETRICAL PACKAGE FOR PC

Jiří Militký
Department of Textile Materials, Technical University
461 17 Liberec , Hálkova Street 6, Czech Republic
Milan Meloun
Department of Analytical Chemistry, Technical University
532 10 Pardubice, Čs. legii 565, Czech Republic
Karel Kupka
TriloByte, Ltd.,
530 02 Pardubice, U Sokolovny 21, Czech Republic

1. Introdution

Chemometrics is a relatively young discipline ranging over chemistry, mathematical statistics and informatics. Its important part comprises methods for the extraction of relevant information from chemical experiments. These methods make use of selected computer algorithms of mathematical statistics, which are discussed e. g., in Meloun *et al.* (1992) and Militký (1989). Some chemometrical computations can be performed by general statistical packages, namely SYSTAT, SAS, STATGRAPHICS. Extensive tests and studies of these and other packages have shown, however, that in many cases they do not support a researcher with suitable statistical methods nor are the numerical algorithms reliable enough (see Militký, 1990).

For these reasons, a new software package ADSTAT has been developed by the authors, who have a long-term experience in chemometrical data processing. ADSTAT consists of relatively independent modules each of which contains several programs or methods concerning one problem. Modules use a spreadsheet-like data editor with some special functions, and a graphical output. The program package is written in Turbo Pascal and makes use of the wide range of possibilities of IBM PC computers. ADSTAT version 2.0 has now started to be marketed by TriloByte, Ltd.

2. Structure of ADSTAT

ADSTAT requires 1 MB of RAM and 8.5 MB on hard disk. EGA and

VGA graphics are supported, as well as printing on matrix and laser printers. Output is provided in ASCII and TIFF formats. The presence of a mathematics coprocessor remarkably increases data processing speed. The program allows the user to work interactively or to analyze data files in batch operating mode.

The operation is simplified by a sophisticated data editor, disk file manager and dynamic graphic output together with general tabular output. Graphic output can be printed in presentation quality on various types of printers or stored into TIFF graphic format files for easy export into other program packages and text editors. ADSTAT uses extensively text windows, pull-down menus and interactive panels. The rules of their use are simple, direct and generally known from other programs. ADSTAT is controlled with keyboard and/or mouse.

Required settings are stored in file ADSTAT.CFG in the current directory. If this file is found, ADSTAT reads all settings and suggests them as implicit values. If this file is absent, ADSTAT creates a new one with start-up implicit settings. Configuration file *.CFG (implicitly ADSTAT.CFG) may be saved at any time. The following settings are saved in a configuration file:

1. Printer type and printing method, size, resolution, etc.;
2. Computing conditions separately for all methods, name and setup of text result file and graph selection setup.

2.1 *Data input*

The ADSTAT data editor is designed to make data input for statistical processing as easy as possible, to enable data file preprocessing and to allow data export and import in ASCII format (with .TXT implicit extension). The data editor uses three value types, which may be contained in a cell: numerical (-1E+99 through 1E+99), text (up to 8 characters) and empty mark.

The editor window contains work area, the local menu and the status line at the bottom with information about editor status, current block, cursor location, etc.

The work area consists of cells of a fixed size. Each cell has its own line and column coordinate.

Data can be entered in the editor in two ways:

(1) by reading data from disk in ADSTAT type file (*.ADD implicit extension) or in ASCII format;
(2) by keying-in the data directly from keyboard.

Data are saved in ADD format implicitly in compact binary code simultaneously with all information about the file, e.g. block definition, variable names. Editor capacity is 30.000 × 255 cells. Real effective capacity depends on memory and disk size.

Besides common character, word, row and block-oriented operations (such as delete, move, copy), it is possible to transpose a block, fill it with numbers with a constant increment and transform data according to a user function. Data can be stored to an ASCII-file on the disk.

2.2 *Program outputs*

All information on processing and results are written to a special result file during computations. This results file is stored in ASCII format, i.e. it is possible for the results to be displayed, printed out and processed with any text editor, as well. Results are arranged into logical paragraphs.

Nearly all methods are followed by graphic output. For some methods (exploratory analysis, residuals analysis) the graphic output is the central point of applied technique and has high confirmative value.

Graphs generated with the ADSTAT program are solely bivariate, because these are more commonly used and their interpretation is easier. Graphical information are saved into temporary file AD{GR}.BIN in internal binary form during the computation. After termination of the computation, the graphs are generated from this file with the Graphs item in menu. This file is also used for graphics in batch mode. It is possible to display up to 10 graphs at a time.

Graphs are displayed on the graphic screen. The graphics screen consists of a few components.

At the top, there is a bar with global icons. The largest area of the screen is occupied by windows with graphs. Each window consist of a frame, having a line with a graph name and graph local icons at the header, and graph body, with axes, curves and points, respectively.

At the lower status line, the most important hot keys for keyboard control are displayed.

The window with icons and graphs is the basic object on the screen. One of the windows is always active. The window can be reduced to the minimal size (iconize). It is then possible only to move it or restore it to its original size.

Graphs may be printed on a printer or written into file in commonly used TIFF black and white graphic format. The second feature is very useful for using the graphic output as report illustration, because most text editors and DTP software accept this format and can incorporate it into text.

2.3 *Organization of computation*

ADSTAT is a multiple windows system managed by a combination of pull-down menu and panels.

Windows may contain a horizontal local menu. Once activated, the windows may be reentered using their order numbers in the upper-left corner.

Interactive panels are controlled with the Tab key first of all, which moves the cursor among individual panel fields. Field items are selected with cursor arrows (the dot in parentheses is relocated in the selected item) or filling in values. Panels contain push-buttons allowing to save setup, recall setup, etc. Some panels contain a field with optional items or parameters, which may be tagged (a cross in a bracket [x]) to activate the item. Some items may have a vertical arrow pointing down one side. The arrow indicates it is possible to choose from a series of prechosen variants (e.g. selection of one of previously specified nonlinear regression models).

The program main menu is located on the first line of panel. The bottom line, or status line, is reserved for current information about the program run, brief help, error reports, etc. The whole area between the main menu and status line is reserved for windows, pull down menus and panels, where all the dialog with ADSTAT takes place. Context-sensitive help containing information about activated items can be recalled.

3. Modules of ADSTAT

In this part, brief outlines of particular modules are given. Attention is paid to the more interesting algorithms and techniques.

3.1 *Basic statistics*

This module includes programs suitable for all phases of one-dimensional sample analysis, comparison of two samples and determination of errors of an indirect measurment. Programs for both discrete and continuous exploratory analysis have special algorithms for detecting suspicious behaviour of data. Another program is used for comparing the sample distribution with ten theoretical distributions on the basis of the Q-Q and TDF graphs (Meloun *et al.*, 1992). Power transformations for an improvement of sample distribution can be provided by using special procedures (Meloun *et al.*, 1992).

In a routine data analysis only basic assumptions are usually required to be tested. Interval estimates of parameters or variances are computed by classical, robust and adaptive techniques.

One program will statistically compare two data samples. Besides classical tests, also more realistic adaptive and robust ones are available. For statistical characteristics of indirect measurements the Taylor series two-points technique and a simulation technique are used.

3.2 *Multivariate data*

In this module, programs are oriented towards processing multivariate chemometrical data i.e., estimation of parameters and their testing on the basis of one or two multidimensional samples. In a special program there is a choice of techniques for graphical representation of such data. Both symbolic and projection types of graph can be used. In the module there are also programs for basic statistical analysis of multivariate samples, correlation analysis and comparison of two multivariate samples.

3.3 *Calibration*

Calibration tasks have usually two phases. In the first one, a calibration model $f(x,a)$ itself is constructed using experimental data. In the second phase, the calibration model (or its inverse) is utilized to determine required response values with a confidence interval. In the case of a linear calibration model, both phases can be simply managed. For special cases of non-linear systems, with a complicated shape of the calibration curve, splines regression models can be successfully used. Another program in this module includes some ways to apply linear regression splines with fixed knots, which can be set up by the user according to three criteria. Quadratic and cubic splines are provided as well.

3.4 *Regression analysis*

This module serves to create linear and non-linear regression models, estimate their parameters and analyze them statistically. For linear regression, special algorithms have been implemented, and the least squares method is only a special case among a series of biased parameter estimations, controlled by a single variable. An analysis of linear and linearized models can be done. Before computations, the data can be transformed to polynomial form, by the Taylor expansion (up to quadratic terms) or generally (any variable is transformed by a user function). A powerfull SVD based algorithm is used (Militký, 1989). A variety of regression characteristics including partial regression graphs can be computed. For diagnostic purposes, a program with over 40 graphs for testing assumptions about the data model and least squares criterion has been included (Militký, 1989).

For non-linear models, an efficient algorithm MINOPT (Militký, 1989). has been developed for parameter estimation and statistical analysis. The regression model is written interactively in the usual algebraic form, with up to nine parameters. Some of the parameters can be set to constant, computation can be interrupted and restarted or aborted in any iteration.

3.5 *Smoothing*

Programs of this group are used for smoothing of experimental data with consecutive derivative or integral analysis. Two cubic spline smoothers are available, those of Reinsch and Späth (Meloun *et al.*, 1992). They differ in smoothing degree parameter and the effect of statistical weights given by the user. The algorithm of Savitzki-Golay is used for simple smoothing based on moving parabolas with adjustable length. Finally, data equidistantly distributed with respect to X-axis can be filtered by various types of digital filters.

3.6 *Analysis of variance*

Programs contained in this module provide one and two way ANOVA. All programs provide, besides a construction of a suitable ANOVA model, also verification of the assumptions on which the model has been constructed, and identification of influential points. One program performs one-way analyses for models with fixed effects and an unbalanced design. Two-way analyses can be performed using another two programs, for fixed and random effects.

A special program is designed for the two-way analysis of unrepeated measurements. Both the classical parameter estimation method and a robust one (median-polish) are used.

The programs also provide tests on the additivity of effects and a power transformation to improve the distribution of the data.

3.7 *Probability models*

In the program of this module, a 'maximum likelihood' method is used to estimate parameters of location, scale and shape for a chosen probability model. Parameter estimation and statistical analysis is realized for normal and two or three parameter Gamma, exponential, Weibull and lognormal distributions.

3.8 *Growth curves*

The growth curve module aims to assist the creation of the optimal type of growth curve and to estimate its parameters including statistical analysis. Nonlinear regression based on least squares is used for parameter estimation. The family of Schnutte's four-parameter growth curve models are offered, as well as logistic, Gompertz's, Mitzerlich's and Richards's models.

3.9 *Chemometric models*

This module contains basic non-linear chemometric models for calculation of dissociation and formation constants, processing of data from potentiometric titration and spectrofotometry. Programs make use of the optimization algorithm MINOPT (see regression analysis). Special programs are devoted to Debye-Hückel model and nonisothermal kinetics. Other chemometrical applications are constantly being added to this module.

4. Conclusion

The program package ADSTAT is built as an open system. Other modules will be added especially for specific chemometrical applications (principal components analysis, cluster analysis, PLS, etc.).

Bibliography

Meloun M., Militký J. and Forina M. (1992), *Chemometrics in Instrumental Laboratory*, Part 1, 2, Ellis Horwood Chichester.

Militký J. (1989), *Mathematical Model Building*, Mimeo, Ostrava.

Militký J., Kupka K. and Rudišar L. (1990), Proceedings Conf. COMPSTAT'90, Dubrovnik.

A COMPUTER PROJECT FOR TEACHING STATISTICS TO TECHNICIANS AND FACTORY WORKERS

Annie M. Morin
ENSAE, E04, 3, Avenue Pierre Larousse
F92241 Malakoff Cedex, France

1. Introdution

Most of the papers dealing with statistical education are concerned either with the training of statisticians, the training of teachers who teach statistics or the training of university students. Actually, statistical education in a school or university environment is easy if you compare it to what happens when you have to deal with technicians or factory workers. On one hand, some of these workers have a low educational level and, even when this is not the case, they left the school system several years previously. On the other hand, they are motivated and remain motivated as long as the teaching is not too formal.

We are presently working on a computer project in elementary statistics for technicians and workers in the industry. The project is financed by a group of firms which wants to give a basic statistical education to their employees. These employees are supposed to have reached a 9th year level. And the goal of the training is to provide them with statistical concepts and methodology that will help them first in collecting data and then in making decisions based upon information. The contents include a presentation of the different ways of gathering data, descriptive statistics with one, two or three variables, time series, quality control and a short introduction to the normal distribution. The product must be self-sufficient. Therefore we have to provide any user with information on the computer without the use of an extra manual.

It is obvious that we have to start with examples to introduce statistical thinking and concepts. Although it is a very usual way to proceed in other countries, it is still rare in France. And we have to avoid the mathematical symbols or to introduce them very carefully. It is not so obvious but it is very interesting for the scientists working on the project. Besides, the material is supposed to be distributed in different firms and factories which have neither the most recent computers nor the most recent software. So, we are developing a product which must work on a common PC. It is not totally finished yet but we plan to achieve it very soon and to test it.

We will present the main steps of our work, the points of interest, and the difficulties we meet in making the project.

2. Presentation of the project

Several companies (car factory, manufacturing industries, banks) have declared that technicians and factory workers must have instruction in elementary statistics. They currently draw Pareto diagrams, use control charts and take a hand on some planning of experiments, without basic knowledge in elementary statistics. Particularly, they are not aware of how important the gathering of data and the quality of collected data are. The companies want a computer assisted learning program which allows in-service training.

The students have reached a 9th year level but some of them did so many years ago. This program, which is financed by a group of firms, must be attractive enough to keep the students until the last chapter. It is difficult to precisely define the audiences. The course must be appropriate for technicians but also for engineers. We have decided that the training is primarily intended for technicians and workers with little statistical education.

Apart the software, we are supposed to provide a user's guide and a battery of tests. The project began in september 1992 and must be achieved at the end of 1993.

3. About the contents

The first chapter deals with the presentation of statistics. We use several examples found in the newspapers and comment on them while introducing the statistical vocabulary. Then the second chapter is concerned with data organization. The third chapter deals with the summarization of data, the fourth one with time series. In the last one, we present the Gaussian distribution and some applications to quality control.

A computer assisted learning program is not a book and we are very careful not to imitate the books. At the beginning of each chapter, a menu is provided. Each idea is developped on one or two screens and, after presenting a notion and an example, we ask the student to do it again. A help on-line is provided, as is a dictionary. A student can access any paragraph of any chapter but if some prior knowledge is required, there is a warning on the screen.

Most of the time, we avoid the use of mathematical formulae. But to deal with the mean or the correlation, it is better to have a formula. The student has access to an icon "maths" where all the formulae are described.

We provide multiple choice items at the end of each paragraph.

4. About the software

The product will be used on microcomputers and "WINDOWS" is required. It has been created with Authorware and doesn't need any statistical sofware. Therefore, students can not process their own data with this program. It was a choice at the beginning of the project (cost, storage memory). We will provide a list of statistical packages that the firm can buy if they want to process their own data (which will probably be the case).

There are several steps in the realization of the computer-assisted learning program:

1 - definition of what is wanted;
2 - scenario;
3 - realization of a prototype;
4 - diffusion.

At each step, there is an evaluation by experts and/or by users. The permanent team working on the project includes a statistician, two computer scientists (the first one manages the project, the second one writes the code), and a graphist.

We have achieved the scenario and work now on the realization of a prototype. We can distinguish two kinds of difficulties:

1 - the first one depends on the audience: we have to avoid formal representations as long as we can and to find a tradeoff between a too literary text and a too mathematical one. We never forget that on one hand the audience does not like mathematics (and statistics is considered as a part of mathematics) but on the other hand they are willing to learn statistics;

2 - the second one depends on the form of training. We write small sequences on a subject keeping in mind that it is not a book and must be more dynamic and alive than a book.

5. Conclusion

Our product is intended for people without statistical education. It is a first level of education but an essential one. We choose daily life examples and our data come from the newspaper or from the firms.

There already exists several products. Most of them have been made for a more highly qualified audience. We insist in our course on data collection. We want to create awareness about data collection among the students. We want to train people to be critical about statistical information they can find in the newspaper and to recognize that, in spite of some bad use, statistics is a valuable tool for their daily life at work or outside work.

TEACHING DATA ANALYSIS WITH THE HELP OF SURVO

Juha Puranen
Department of Statistics, PL33 (Aleksanterinkatu 7)
00014 University of Helsinki, Finland

1. Survo

The Department of Statistics at the University of Helsinki uses various statistical software packages such as Survo, SAS, SPSS, SPLUS, SAZAM, PC-GIVE, GLIM, EGRET... Among these, the local statistical software package SURVO, created by professor Seppo Mustonen (Mustonen, 1992), is the most widely used.

Outstanding in Survo is the editorial user interface in which all the functions are carried out. Outcomes from various statistical analyses are also recorded in the edit field.

Using Survo's editorial capabilities, it is possible to generate reports, incorporate Survo-generated graphics, and send the overall report to the printer. For example the 500 page Survo User's Guide has been created in this way, saved onto disk, and taken to the printer as a Postscript file.

Comparing Survo to other statistical software, one can see the resemblance to S-PLUS, when this latter one is used within the GNU-Emacs environment. The number of different statistical techniques is slightly higher in S-PLUS. However, the user interface in Survo is much better.

Students can get 'SURVOS', a freeware version of Survo, for their personal use. Although the number of statistical procedures in 'SURVOS' is limited, it is far more complete than most existing freeware programs.

2. Survo in teaching

During the last 9 years I have used Survo's capabilities for almost all my lectures. By means of Survo I have:

handled all my text processing;
created transparencies;
planned and solved exercises;
built up all required teaching material.

3. Survo in data analysis

In this paper I will present the use of Survo for teaching data analysis where emphasis is focused on the handling of statistical data, the selection of the right statistical method and the interpretation of the final results.

The background of the students attending the course is usually merely the introductory course (40 hrs lectures + 20 hrs workshops) and the supplementary course (40 hrs lectures + 20 hrs workshops) in statistics. Almost 95% of the students have at least basic knowledge of computers but hardly any knowledge of real statistical data analysis.

The course on data-analysis is put together from three parts:

1 - Lectures 20 hrs (theory);
2 - Survo laboratory 15 hrs;
3 - Individual case study.

During the lectures, we work through various data-analytical problems, such as how to describe data, what statistical method to use, and how to interpret the results.

During the Survo laboratory, students learn how to analyse data in practice. All the teaching takes place in the computer laboratory, equipped with eight 486-PCs and the teacher's personal PC connected to a Data-show projector. For the laboratory the students are divided into small groups of 8 to 16 so that there is no more than 1 or 2 students sharing one computer.

Only a few students have their own data sets for a genuine case study. For the majority of students a common data set is selected. Last year's data set dealt with people's support of Finish political parties. For the case study, each student selected a personal random sample of 120 rural and urban districts. The main object of the case study is to explore the distributions of support for the parties they selected. This support, of course, depends on various confounding variables (demographic, socioeconomic, etc.).

Some students prefer to use SAS or SPSS to carry out their case study and therefore do not participate in the Survo laboratory.

4. How to use Survo in teaching data analysis

4.1 *Lectures*

A few years ago I used transparencies in all my teaching. "God knows how many times the right transparency was seemingly missing in the disordered pile of transparencies". With the versatility of Survo, I was able to start a computerized and more disciplined teaching presentation.

A teaching presentation is usually about 5 to 10 minutes long. It is built up from texts, graphics and simulations. Because the length of each teaching presentation is short, it is very easy to pause, start it again, or reorganise it.

For the presentation of the teaching material a portable computer and color Data-show projector are used. In some special situations, extra transparencies or a blackboard might be used.

4.2 *Survo laboratory*

In the Survo laboratory, students learn how to analyse data in practical situations and how to use the statistical package Survo.

Usually the students have a very limited experience in statistical computing. Therefore, teaching should concentrate on both the use of the Survo package and data analysis methodology. Since the Survo laboratory duration is limited to only fifteen hours, teaching focuses on both simultaneously.

The Survo laboratory includes the following topics:

1 - Text processing and data input;
2 - Data file management;
3 - Data aggregation and graphical presentations;
4 - Hypothesis testing;
5 - Report printing.

A collection of short tasks (30 in all) is used for teaching purposes. The tasks are chosen so that they support as much as possible the final case study.

The Survo laboratory is intensive. There are more tasks than the students can perform under normal circumstances. In this way both advanced and less advanced students are kept under constant pressure! At the end of each session about 1 or 2 tasks might remain undone for some students. These are, however, to be completed before the next workshop.

Students learn the use of Survo as a by-product in the process of solving their practical assignments. In this way, emphasis is directed entirely to data analysis by means of Survo. Experience shows that students are more motivated to learn about the software Survo, when they are faced with real research situations.

In order to be able to control heterogeneous groups of students with different levels of skill, two teachers follow the students' progress. A principal teacher shows how to carry out analysis and how to interpret the results. The second teacher is either an assistant professor or an advanced student who watches constantly the progress of all the students in their problem solving attempts.

Because students have a limited knowledge of Survo, they can get stuck in more or less simple matters. In such a case, they can get instant advice either from the main teacher or his assistant. This allows them to concentrate on analysing the data and, therefore, avoid conflict with the system they use.

4.3 *Case study*

The case study analysis starts immediately after the laboratory is over. Since students' ability to analyse data at this stage is rather uncertain, a help desk is organised where they can get thorough advice whenever necessary.

5. Why Survo?

Survo is an excellent tool for this type of teaching. Survo's unique user interface makes possible the control of text, data, graphics and statistical analysis.

Also the use of live simulations and the alteration or addition of a single or a multiple set of observations give a dynamic aspect and a new dimension to students'comprehension of the course. This undoubtedly motivates the students more towards learning than does the traditional ready-made examples method.

By means of Survo the students can perform their analyses step by step and learn more about statistical methods. Survo is not a 'black box' where one inputs data from one end and gets the outcomes from the other end. In fact, a knowledge-based procedure is required from any prospective student in order to get the best results using Survo.

By means of the editorial user interface, students can make personal notes in the edit field. They can save extra information made up of 'whats and whys' they might use later on. They also write into the edit field what methods were used and how the results were interpreted. The resulting document can be saved on disk and/or printed out and used later on as a model when analysing the case study.

6. Future

In the future, the improvement of the teaching material concerning the theoretical part of the course will continue. In addition, I am planning to install the latest version of the teaching material in the local computer network. In this way, students will be able to access the teaching material through LAN and make this part of the course without a need to be present during the main lectures.

Bibliography

Mustonen S. (1992), SURVO *An Integrated Environment for Statistical Computing and Related Areas*, Survo Systems Ltd, Helsinki.

EDUCATION PROGRAMMES AND TRAINING IN STATISTICS

Invited papers	Vic Barnett James P.M. Ntozi *et al.* Rudolf Teekens
Contributed papers	Shrikant I. Bangdiwala Andrés Carrión García A.B. Raha Ante Rozga Susan Anne Starkings

THE ROLE OF CONSULTANCY IN UNIVERSITY EDUCATION AND PROFESSIONAL TRAINING IN STATISTICS

Vic Barnett
AFRC Institute of Arable Crops Research
Rothamsted Experimental Station
Statistics Department, Harpenden Herts AL5 2JQ, UK

1. Introduction

Statistical education and training at tertiary level encompasses a very wide range of activities. There are the **short "service" courses** for students studying, say, engineering or medicine or politics (often of just 10 or 20 hours providing familiarisation with basic concepts or instruction in particular techniques e.g. testing or regression) often nowadays linked to some computational software system. At the next level we encounter in some disciplines such as agriculture, economics or social sciences, a number of "**designer" courses** spanning a fair proportion of the study time, perhaps up to one third of the overall programme. Finally, there are **full statistics programmes** for students of statistics *per se* (or mathematics programmes oriented to non-deterministic mathematics) where most of the study time is spent on the concepts, methods and practice of probability and statistics, with the inevitable supporting coverage of relevant mathematical material.

We shall be concerned in this paper specifically with the education and training of *professional statisticians*: those who will make their career in developing statistical science and/or solving statistical problems on a day to day basis. These are not likely to be found in our first category above. The second category will yield some, but in highly specific form perhaps as *biometricians, econometricians or psychometricians*. The majority of the professional statistical fraternity particularly those who will ply their trade and do methodological (or even conceptual) research across the disciplines will come from the third category. Here is the dilemma! There is a fundamental conflict of aim (or at least of interpretation) in the statistical education and training of this group, which usually takes place through the avenue of the universities.

What is this conflict? All would agree that it is important to encompass the variety of aspects of the subject:

(i) *concepts and principles*
(ii) *methodology*
(iii) *applications and practice*

and that each requires appropriate in-depth coverage. The conflict arises in the different interpretations of what is meant by "in-depth" and "coverage". Consider what happens in practice. Different departments are not all large enough, nor necessarily have the capability (or even the desire or time) to meet all the needs. Some may express through their teaching programmes quite different interpretations. Let us note some examples of this.

(a) In presenting *concepts* and *principles* (and indeed methodology) there are differences of view on what are the appropriate background skills, e.g. in mathematics or computation. For some these involve detailed calculus, algebra, systems analysis, etc. at a high honours level. ("How can we start to deal with stochastic processes without extensive complex variable theory, or software engineering?"). For others a modest understanding of linear algebra and use of a few software systems may be felt to be adequate.

(b) Wide-ranging exposure to *statistical methods* is crucial for the professional statistician. Does this mean statistical tests, confidence intervals and a few linear model methods; or does it extend to time series, multivariate analysis, nonparametric methods, survey sampling, spatial modelling, design of experiments, robust methods, Bayesian analysis, design theory, programming, specific statistical software systems, etc?

(c) Coverage of "*applications*" is essential, and widespread *practice* of the techniques and methods is vital (otherwise it is akin to training an airline pilot on a flight simulator without ever letting the trainee in the cockpit of a plane). It is here that the greatest differences are to be found in what is regarded as appropriate "practical" experience. On some courses only occasional sterilised "practicals" are given. For example a one-way analysis of variance is explained and some numbers given on which to work out the sums of squares and to do an F-test. And "applications" might consist of remarking that some particular methods tend to be used in industry and medicine. Of course this is a parody! But how many programmes of study really stress the applications in real depth, have a comprehensive programme of practical case-study work from the beginning to the end, and have a structured professional approach to training the student to become a skilled statistician for various workplaces (whether these are the pharmaceutical industry, the finance sector, the research bench or the university)?

Do I ask too much? Perhaps so! There is space for different aims and emphases and all-too-little time to achieve any of them. Such is the widely expressed response. This is not really good enough, however, and the big-wide-world is ever critical and demanding. One consideration is paramount and cannot be sidelined. Whatever the subsequent destination of the statistician, an ability to *use his knowledge* is vital for judicious choice of methods as are *research skills* for the inevitable demands of constantly needing to develop the subject for the practical environment. We should never forget that what we teach is essentially illustrative, not all-embracing. "Who has ever carried out a *t*-test on normal data, knowing that the data were in fact normal, and that such a basic procedure was fully adequate for the problem in hand?".

Of these three areas of conflict it is the question of what is the appropriate in-depth coverage of *practical statistical and applications* that is the most important, and where the greatest failure arises in many education and training programmes for statisticians.

We shall consider what the real needs are in this respect how they are sometimes not met and how the situation can be improved, particularly through close contact with consulting work. The general principles will be illustrated with examples from consulting experience.

2. The practical emphasis

About 20 years ago, from a university stance, I remarked: "a conflict ... can sometimes arise between the commercial pressures which demand a quick and easy answer and the academic aim of a full study and formal solution. We can detect some reluctance in the profession to 'soil hands' with messy real-life problems" (Barnett, 1976).

Nonetheless there is no doubt in my mind that the function of statistics is to solve real problems (across all subjects from agriculture, through medicine to zoology). The stimulus to research (however mathematical it may become) should be a practical one: the end product of statistical analysis is to clarify understanding of, or the taking of decisions in, down-to-earth problems involving indeterminism.

Let us consider an example where the premium was on a simple and quick answer.

- Barnett and Hilditch (1993) describe such a problem in the context of *the* *use of artificial pitch* (AP) *surfaces in English football.* The Football League was preparing a report (Football League, 1989) giving its future policy on whether APs should continue to be used in the four Divisions of the English Football League. The concern was that APs might confer an unfair advantage on the home team. To determine this they sought statistical evidence, requiring a *rapid response* from *limited data* on an *issue which was poorly reflected* in the database (very few teams had APs).

These were typical consulting difficulties but make good "case-law" for teaching practical statistics. Thus in the education and training of statisticians a central component must be instruction on how to answer real-life practical problems – to a deadline – across different disciplines – and with clearly understood answers.

The only data available in the football problem were the end of season results (points, outcomes, goals) for the four divisions over an eight-year period during which only 4 out of more than 80 teams used an AP (and then not for all the time). In fact only 362 out of 32468 games were played on an AP. Could we hope to detect an effect in such sparse data?

Results can be expressed in terms of *points* (3 for a win; 1 for a draw), *outcomes* (win, draw, lose) or *goals* (scored for each team in a match). These provide alternative (but not independent) measures of relative performance. Consider just *points.* The points scored at home and away for teams with home APs or NPs (natural grass pitches) were as follows (percentages of points scored at home and away are in parentheses):

	Home	*Away*
AP	718 (66.1)	368 (33.9)
NP	10860 (62.6)	6476 (37.4)

So we see that AP teams score a higher proportion of their points at home compared with NP teams (0.661 and 0.626, significantly different at 5%). Does this mean that using an AP confers an advantage? This is not immediately clear. Suppose *better* teams tended to score a lower proportion of their points at home, and AP teams tended to be *poorer* teams. We would expect to see the sort of effect observed without it necessarily reflecting in any way on use of AP. In fact the proportion of points scored at home $p_H / (p_H + p_A)$ did depend on team quality: see Fig. 1 (reproduced with permission from Barnett and Hilditch, 1993). It was also noted that the AP teams did not seem to cluster at the top or bottom of the division, so that such dependence might therefore not be too critical. However the dependence (between end-of-season league position and proportion of points scored at home) sounded a cautionary note and it was felt necessary to seek measures of relative performance (H *versus* A), which were *not related to the intrinsic quality of the teams.*

This was studied for all three ways of expressing performance: points, outcomes, goals. Such positionally independent relative measures were found in each case. For points, the measure took a particularly simple form, $p_H - p_A$. The overall results in terms of average values of this measure yielded 1.50 for the AP group and 1.36 for the NP group, a highly significant difference.

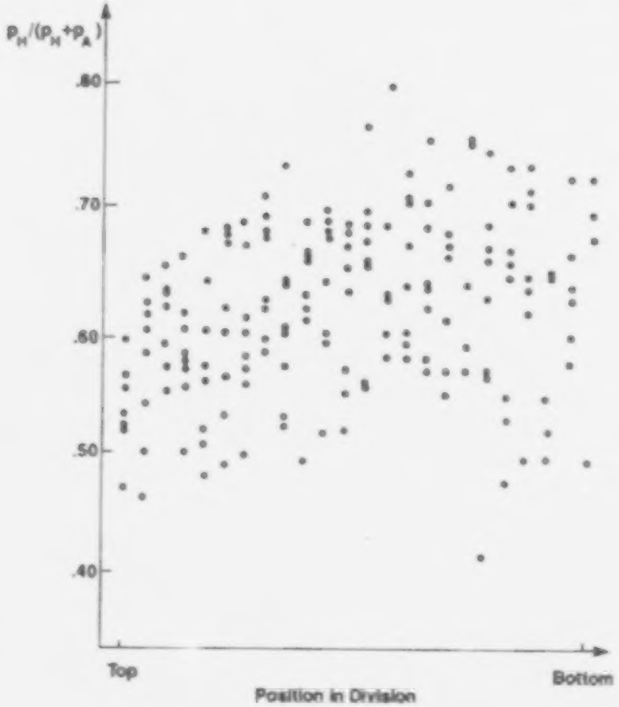


Figure 1. *Points ratio and end-of-season positions (1987/88 and 1988/89)*

We could be much more confident that we were seeing differential performance more related to pitch effect (in that it is no longer confounded with team quality). It was not a trivial difference: the AP seemed to be worth about an extra 5-6 points (or 2 wins) per season.

3. Teaching practical statistics

The teaching *of practical statistics* is not straightforward: it is accordingly often played-down. Sometimes this is from genuine belief that fundamental philosophical, conceptual and mathematical components are paramount. More often it arises (I believe) from the lack of experience or confidence of the lecturer, who has probably never worked as a practical statistician or even faced a client across the consulting table.

Many have written on how to teach practical statistics in particular contexts (see for example: Bishop, 1964; Deane, 1964; Finney, 1968; Greenfield, 1979; Spicer, 1964). Anderson and Loynes (1987) present a detailed general study of this topic, basing it on a unique and powerful type of practical course that we pioneered in Sheffield University. They too start from the premise:

> "Statistics as a subject is inseparable from its application and practice .... any statistical education should take account of this connection".

Anderson and Loynes (1987) identify several components of practical training. These include:

- *drill exercises* (rather obvious methodology applied to simplified datasets with only limited reference to the practical environment: these must have a limited role to play);
- *statistical experiments* (practical problems set up in the teaching environment leading to data collection, analysis and report: important but time-consuming and necessarily rather remote from real-life in order to be manageable);
- *critical reading* (of well selected practical publications: a useful secondhand experience);
- *projects* (varying in size from 2 weeks to several months: vital study of extensive datasets with much background practical detail and requirements for written, and possibly oral, reports);
- *outside contacts* (meetings with applied statisticians to hear from the workbench how it should be done! This could usefully include sitting-in on consulting sessions, and hearing detailed presentations of case-studies);

placements (in different applied fields and for long enough to gain valuable experience, but who manages to find the time for this?).

In all these contexts action and interaction are essential, with copious opportunity to work in a group, be involved in discussions, and make presentations of practical work. There is one underlying aim: to teach the student how to *communicate* with the research worker, often on matters of social concern. "The statistician gains in experience and ability to employ his art by contact with diverse disciplines. This also stimulates new developments within the armoury of statistical methods. The research worker gains in the realisation of what Statistics has to offer for the understanding of his own subject area. Thus this role of the statistician as a consultant and collaborator has vital educative overtones" (Barnett, 1983). Many others have written on this, e.g. Bartholomew (1973), Benjamin (1971), Sprent (1970), Chanter (1978), Freeman (1978), Greenfield (1979), Marquardt (1979), Hooke (1980).

An example of social-impact consulting with much potential for developing cogent teaching material recently arose in the field of *false alarms from security systems*. It is well-known that the false-alarm rate is high (whether in intruder alarms in industry, shops or the home). But how high? The false-alarm rate can be as high as 85%-90% overall, with the contingent vast waste of natural resource due to security companies or the police making unnecessary responses.

The conventional way of trying to deal with this problem is to seek to make alarm systems more and more reliable and sophisticated (and inevitably more expensive). One large security company posed the question in *statistical*, rather than *technical*, terms: was it possible to obtain a more accurate assessment of whether an alarm call was genuine by taking account of a variety of concomitant information such as day, time of day, type of property, weather conditions, nature of alarm system, and so on? The argument was that if an alarm call from a corner shop at 3 am on a clear Sunday morning was ten times as likely to be genuine as an alarm call from a supermarket at 5 pm on a wet Monday, we would want to give the former higher priority for attention than the latter.

To examine such prospects of differing false alarm rates, a major police database of over 100,000 calls over a three-year period was statistically analysed. A *multiparameter log-linear model was fitted* including as many main effects and interactions as were necessary to obtain the maximal significant fit. Sure enough, major influences were found. For example, the estimated genuine alarm rates varied from less than 2% to more than 25% (a twelve-fold difference) at different times of the day (peaking between midnight and 3 am). Different days also produced quite different levels of estimated genuine alarm rates (with Saturday, Sunday, Monday and the remaining days all differing from each other), as did the different types of property (from as low as 2% for private houses to approaching 20% for certain classes of retail and public property), and so on for other factors.

Whilst some of these differences were qualitatively not surprising, the wide discrepancies of value were not previously known. Most importantly, however, major interactions were found even up to the 4-factor level. Consider as an example of first order interaction the estimated genuine alarm rates (in percentages) for retail properties at different times of different days:

Hours:	*02-06*	*06-10*	*10-14*	*14-18*	*18-22*	*22-02*
Saturday	25	1	1	3	6	14
Sunday	36	2	2	7	10	19
Monday	47	2	3	3	8	28
Other days	39	2	2	3	7	22

These differences (from 1% to nearly 50%) are enormous in their relative import and surely provide a vital basis for differential response to alarm calls depending on time, day (and other factors). They show high potential for a *source-sensitive approach to response* which could save much of the unnecessary expense of an undifferentiated response system.

Of course there are problems: social attitudes to police and security-firm policy, geographic effects, and so on. But the real potential is demonstrated, for a diagnostic computerised software system to aid responsible attention to calls for help! And what a rich field for practical illustration of statistics in the teaching process.

Thus, returning to our main theme, we conclude that consultancy is central to the role of the statistician. It is obvious that we should seek *to teach consultancy* (see Cox, 1968; Rustagi and Wolfe, 1982; Zahn, 1982) as an integral part of statistical education (even if not all are agreed on how to, or are personally equipped to be able to, do it).

What is less obvious but even more important is the converse: the *role of consultancy in teaching practical statistics.* Our two examples already illustrate the possibilities.

Let us consider in more detail what this means!

4. Consultancy as a basis for statistical education

Consider again the main problem. How do we instill an ability to *actually* *do statistics* rather than just *knowing in principle how we might do statistics.* This includes being able to provide confident, rapid response; knowing that sometimes sophisticated methodology may be needed, whilst on other occasions it may not be and ingenious data-dexterity will yield the answer; recognising that basic research is required in some problems but that a "quick and easy" answer should be offered as a stop-gap; developing the manner and skills to impress others with the fact that we understand their problem and know how to provide an answer. Another recent consulting experience completes the spectrum: from *simple data-analysis* (e.g. the football problem), through *formal methods* (the alarm systems example) to *fundamental research* (in the example we now consider).

The problem arose in the public domain in seeking methods of estimating road-accident rates at junctions from data on traffic flows and the physical geometry of the road junctions. The work is described in Barnett and Wright (1992), and is shown to have applications in many other fields (e.g. medicine and biology). It looked initially to be the fairly standard regression type of investigation so often encountered in consultancy situations, but it turned out on detailed consideration to be much more complicated: in fact to be a previously unresolved fundamental research issue.

Essentially we had observations of a discrete random variable Y (long-term accident rates) of Poisson form with mean value function $\mu(\mathbf{x})$ depending on a variety of predictor variables $\mathbf{x}$, some of which were also discrete counts. The specific form of $\mu(\mathbf{x})$ was

$$\mu(\mathbf{x}) = \exp(\alpha'\mathbf{x})$$

where α is a set of parameters. Thus it had the characteristics of a generalised linear model (GLM) with logarithmic link function and Poisson error structure. However, there was a "sting in the tail". The predictor variables were in two distinct groups $\mathbf{x}_1$ and $\mathbf{x}_2$, with $\mathbf{x}_1$ referring to traffic flow rates and $\mathbf{x}_2$ to physical characteristics of the road junction. The $\mathbf{x}_2$ values were able to be measured very accurately but the flow rates could only be estimated from short-term traffic counts, and were correspondingly inaccurate ("measured with error").

Thus

$$\mu(\mathbf{x}) = \exp(\alpha_1'\mathbf{x}_1 + \alpha_2'\mathbf{x}_2) \tag{1}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are of dimensions r and s (in fact $r = 2$), with $\mathbf{x}_2$ observed without error and $\mathbf{x}_1$ only measurable with error.

This latter property moved us away from the GLM to a model of functional form, so that we had what might be termed a *generalised linear functional (Poisson) model.* Essentially we had an independent random sample $y_1, y_2, \ldots, y_n$ of observations of a Poisson random variable Y with mean (1) subject to the different status of $\mathbf{x}_1$ and $\mathbf{x}_2$. The client required a maximum likelihood solution implementable as a macro in one of the principal statistical software systems.

Here was pressure indeed ("miracles take a little longer!"), but in fact a solution was obtained and a macro produced, all to the required deadline.

Thus we have reached a stage where it is clear that all aspects of statistical education depend vitally on close contact with the consulting world.

The key to how this can be achieved is that the teaching environment itself (the university department, say) *should be fully committed to consultancy even to the extent of running a commercial and comprehensive consultancy service.* Should university statisticians do (industrial, medical, commercial, etc.) consulting? Depending on background qualifications, experience and (particularly) nationality, the immediate answer lies on a wide spectrum, from

Of course not - we do the important (difficult) fundamental work

→ to ←

Of course, how else can we teach, do research, extend the subject?

I am biased (I come from a totally *empirical* tradition in the UK). I automatically regard statistics as designed *to answer real life questions.* From the outset (over 30 years ago) I have always as a university statistician done "consulting" in this sense.

So my answer is "of course"! I earlier observed (Barnett, 1983):

"*The service, and statistical advisory, role, of a university Statistics Department is vital for creating the proper educational environment.* I believe that any Statistics Department must be seen as providing a university-wide practical facility, both in the teaching of students from other disciplines but, more important, in *offering a full statistical advisory service* to all university staff, and postgraduate students (and to outside organisations, if logistically feasible). See Tukey, Box, Hartley and Kempthorne (1968), Gibbons and Freund (1980). This requires more than just a casual attitude of being prepared to help if anyone happens to ask. The advisory service should be widely advertised. Staff (not necessarily all) must be committed to providing the service, and have the confidence to do so.

The university needs to recognise and welcome the facility, to the extent of providing adequate resources, including computing equipment and a staff member to conduct computational work. There is nothing worse than offering a service and saying that someone's problem is soluble in principle, *but* that lack of time, computing power, etc. prevents it from actually being solved".

In advancing the importance of a consulting base in a department of statistics we need to ask:

Why?
How?
What are the benefits?
What are the difficulties?

In recognising that the practical skills have to be passed on to students and trainees, we must also remind ourselves that the "teachers" often do not possess the experience on which to do this, either in terms of what or how. It is in this dilemma that the consultancy base of a department can be so vital. It provides the very experience that may be lacking, day by day, in an exemplary manner.

Specifically in response to the question, *why*?, we might respond:

- for practical case material for teaching
- for consulting experience for teacher and student
- for research stimulus
- for social responsibility (we should solve problems)
- for income generation (*why not?*)

But can we do this: *how?*

- by encouraging every individual to develop and promote external links (with no financial redress)
- by setting up (in due course of time) a formalised advisory or consulting service (even a limited company) with its own staff augmenting the tasks done by those of the general teaching and research staff of the department who wish to be involved.

What are the difficulties? These are readily summarised:

- staff inertia (*personal lack of confidence*)
- official obstacles (*institutional lack of confidence*)
- lack of time (*claimed*)

- legal constraints (*usually resolvable*)
- fear of responding to immediate pressures (*customer wants the answer yesterday!*)
- difficulties (*or lack of know-how*) in drawing up and costing contracts.

Such problems are real, but I do believe that they are more than outweighed by the great benefits that accrue to a statistical education and training programme from operating its own consulting service. These should be obvious and clearly include:

- stimulus to professional development (*individual and departmental*)
- enhancement of reputation and of contact base
- improved research and teaching climate
- financial flexibility.

What is most important is that the modern world will no longer be prepared to accept (as it did in the past) statistical education and training which is not demonstrably being provided to satisfy practical needs. And this must be done by statisticians, themselves well-equipped and well- experienced in carrying out the practical professional performance they claim to be presenting to their students. *Can any one of the statistical educators avoid being a consultant?*

Bibliography

Anderson C.W. and Loynes R.M. (1987), *The Teaching of Practical Statistics*, Wiley, Chichester.

Barnett V. (1976), The Statistician: Jack of all Trades, Master of One? *The Statistician*, 25, pp. 261-279.

Barnett V. (1983), *Statistical Education*, Eusko Jaurlaritza-Gobierno Vasco.

Barnett V. and Hilditch S. (1993), The Effect of an Artificial Pitch Surface on Home Team Performance in Football (Soccer), *Journal of the Royal Statistical Society*, A, 156, pp. 39-50.

Barnett V. and Wright D.E. (1992), Biomedical applications for a generalized linear functional Poisson model, *Journal of Applied Statistics*, 19, pp. 41-47.

Bartholomev D.J. (1973), Post experience training for statisticians, *Journal of the Royal Statistical Society*, A, 133, pp. 65-70.

Benjamin B. (1971), The Statistician and the manager, *Journal of the Royal Statistical Society*, A, 134, pp. 1-13.

Bishop H.E. (1964), The training of government statisticians, *Journal of the Royal Statistical Society*, A, 127, pp. 211-215.

Chanter H.E. (1978), Communication and the statistician, *BIAS*, 5, pp. 49-56.

Cox C.P. (1968), Some observations on the teaching of statistical consultancy, *Biometrics*, 24, pp. 789-802.

Deane M. (1964), The training of statisticians for economics and business, *Journal of the Royal Statistical Society*, A, 127, pp. 216-218.

Finney D.J. (1968), Teaching biometry in the University, *Biometrics*, 24, pp. 1-12.

Freeman G.H. (1978), Consultation and research in a statistics section, *BIAS*, 5, pp. 22-32.

Gibbons J.D. and Freund R.J. (1980), Organisations for statistical consulting at colleges and universities, *American Statistician*, 34, pp. 140-145.

Greenfield A.A. (1979), Statisticians in industrial research: the role and training of an industrial consultant, *The Statistician*, 28, pp. 71-82.

Hooke R. (1980), Getting people to use statistics properly, *American Statistician*, 34, pp. 39-42.

Marquardt D.W. (1979), Statistical consulting in industry, *American Statistician*, 33, pp. 102-107.

Rustagi J.S. and Wolfe D.A. (eds.) (1982), *Teaching of Statistical Consulting*, Academic Press, NY.

Spicer C.C. (1964), The training of medical statisticians, *Journal of the Royal Statistical Society*, A, 127, pp. 219-221.

Sprent P. (1970), Some problems with statistical consultancy (with Discussion), *Journal of the Royal Statistical Society*, A, 133, pp. 139-164.

Tukey J.W., Box G.E.P., Hartley H.O. and Kempthorne O. (1968), The Future of departments of statistics, in Watts D.G. (ed.) (1968), *The Future of Statistics*, Academic Press, NY, pp. 103-137.

Zahn D.A. (1982), Teaching Statistical Consulting: Statistical and Non-Statistical Aspects, in Rustagi and Wolfe (1982).

THREE DECADES OF TRAINING GOVERNMENT STATISTICAL STAFF IN DEVELOPING COUNTRIES: THE AFRICAN EXPERIENCE

James P.M. Ntozi and George W. Kibirige
Institute of Statistics and Applied Economics, Makerere University
P.O. Box 7062, Kampala, Uganda

1. Introduction

Almost all official statistical services in Africa were introduced during the colonial era with the setting up of statistical units in the colonies and territories whose functions were largely determined by the ruling colonial powers. In East Africa, the service was set up in 1926 as the Statistical Section of the East African Governors' Conference with a mandate of "collecting statistics gradually, on the same method, throughout the territories, and to tabulate and compare results so that true inferences can be drawn" (Singh, 1971). Nigeria's statistical office was established in 1947 with personnel deployed from the Treasury, Customs and Office of the Chief Secretary.

Belgium and France set similar units in territories under their control to produce statistics for the administration of the colonies and provide a framework for the orderly conduct of export and import trade.

During the reign of the colonial powers, often there was no impetus for the training of indigenous statistical personnel. Citizens of the administering power provided the small core of professionals to operate the units.

The count-down period to independence in the witnessed 1950s an increasing demand by the countries for urgent data on social and economic conditions. The pressure for expansion of national statistical services was intensified at independence or soon after (1960s and 1970s) as countries sought more comprehensive and reliable data to formulate and implement satisfactory development programs. This expansion, however, was being hampered by a serious death of statistical personnel, not only professionals (qualified statisticians), but also supervisors and assistants needed to carry out censuses and surveys as well as to process data compiled in these and other activities. Efforts were made at supplying statistical personnel to the services through the establishment of local and regional statistical training programs as well as support of formal study at universities, professional training centres and other institutions of higher learning, and in service and on the job training programs.

2. Training government statistical personnel

During the last thirty years a number of training schemes and projects have been devised to assist in the training of personnel for government statistical services and improve on their supply. These schemes can be grouped into four categories: middle level, professional level, postgraduate training and re-training, and area or subject specific short-term courses and seminars.

3. Middle level

Initial training efforts arose out of the pressure for data compilation as statistical needs increased in the countries. The focus was on producing personnel who could assist the few professionals in the country in supervising censuses and surveys and processing of data compiled through such activities. A need therefore arose for a training scheme whose emphasis was more on practical aspects of statistical operations. Regional and international training centres were established for this purpose. The trend of establishing centres at regional and international level was due to shortage of professionals in any one country to carry out the training activities and the cost considerations. These centres have continued to train semiskilled statistical personnel for many countries. One shortcoming which arose with regional centres was that the practical training tended to be characterized by the social and economic condition of the host country which at times was at variance with conditions applicable in other participating countries. The language problem in fieldwork was a major handicap.

Many national statistical offices require a large number of personnel trained at this level. Only a few from each country can be trained at regional centres or abroad. Besides, releasing trainees on a full-time basis for long period is at times not feasible. Some countries established national training centres so as to bring the training content near to the actual requirements for each country.

Besides the setting up of national training centres, countries have organized in service and on the job statistical training programs. However, national training programs at the middle level have faced the problem of getting competent teaching staff. In some cases external lecturers from universities have been used. These, at times, have not been found suitable especially in handling practical areas. There is also a lack of appropriate materials at this level.

4. Professional and postgraduate level

Professional level training was initially given in overseas institutions and universities. African statisticians who manned statistical offices after independence were almost invariably trained in U.K. or France. Later, institutions were established from which professional statisticians could be recruited.

In 1978, the Statistical Training Programme for Africa (STPA) was set up with the primary objective of making Africa to be self-sufficient in the supply of statisticians by the 1990s. A ten-year program was set up by the United Nations Economic Commission for Africa (UNECA) with the assistance of various donors to enable STPA to obtain its objective.

This was to be achieved by increasing the number of trained statisticians and improving and maintaining the quality of the current statistical staff. It was envisaged that during the 1978-87 period, the STPA centres would increase their output by 50%, with greater increases in the professional training of the Anglophone centres. In addition, a special project for training professional statisticians in the Portuguese-speaking African countries was to be implemented. Hence, the centres were to aim at improving their efficiency and output, make their training programmes more practical, and ensure self-sufficiency.

Furthermore, the top priority of STPA was to produce professional statisticians. Such programs aimed at offering a basic university training leading to one of the following degrees (diplomas): ingenieur de travaux statistiques, ingenieur d'application de la statistique for the Francophone countries and first degree in statistics (Bachelor of Statistics or equivalent qualification) for the English-speaking areas.

Related to the professional statistical training were the more specialized courses offered to the statisticians at the postgraduate level. These scholars would then become specialists in various fields, such as national accounts, agricultural statistics, sample surveys, demographic statistics, economic statistics, labour force statistics, social statistics and statistical computing. These programs would lead to qualifications, such as Certificat d'études Supérieures Spécialisées, graduate diplomas in Statistics, or Masters and Doctorate degrees. In addition, short term courses, workshops and seminars were to be organized for working statisticians by the local STPA centres or foreign institutions.

5. Conclusion

A number of training schemes and programs have been devised to ensure a steady supply of statistical personnel to governments statistical services in Africa. However, the statistical needs of the countries have continuously been diversifying and increasing dramatically. The schemes and programs have to adjust accordingly. In addition, the loss rate of trained personnel at various levels is quite high. One of the solutions to this problem would be that of improving the supply position.

Bibliography

Ntozi J.P.M. (1992), Training of African Statisticians at a Professional Level, *Journal of Official Statistics*, 8(4), 1992, pp. 467-479.

Singh D.C. (1971), The East African Statistical Department, *Economic and Statistical Review*, 41, pp. 8-11.

THE EUROSTAT PROJECT "TRAINING OF EUROPEAN STATISTICIANS" (TES): EXPERIENCE OF THE FIRST THREE YEARS AND SOME REFLECTIONS ON THE FUTURE

Rudolf Teekens
"Training of European Statisticians"
Bloc B-6 2-6, route de Trèves, L-2633 Senningerberg, Luxemburg

1. Introduction

Eurostat (the Statistical Office of the European Communities) started the project "Training of European Statisticians" (TES) at the beginning of 1990. The project has the following objectives:

- to promote a convergence of the practices of National Statistical Institutes (NSI) by means of common training measures and by creating a context in which methodical (but unofficial) comparison may be made of national statistical systems and experiences, thus contributing to the harmonisation of concepts and methods;
- to train and/or retrain statisticians in the use of new and efficient operational methods and tools and to explore technological frontiers in the field of applied statistics;
- to offer to young people trained in other disciplines the possibility of specialising in statistics, in a broader context where account is taken of the Community dimension of problems and to present to statisticians the new statistical requirements which have resulted from changes in European society and stimulate reflection on different models.

The above objectives are being pursued through the following activities: the organisation of annual professional training programmes for statisticians; the creation and the maintenance of a data bank on existing post-university training courses in Europe relevant to statisticians, followed later by the issue of a bulletin which will serve as a means for exchange of information on training facilities and the dissemination of information on ad-hoc events (conferences, seminars, etc.).

2. Characteristics of the TES training programme

2.1 *Objectives*

The objective of the programmes is to offer professional training facilities of a strongly applied character on subjects which, by their nature, cannot be dealt with adequately at universities, are specific for the work environment (statistical institutes) or have a dimension surpassing the national boundaries. Moreover, the programme offers a platform for high-level discussions on various management and policy issues relevant to National Statistical Institutes. The programme is to be seen as complementary to existing national professional training facilities. It should be stressed that the programme aims at training and development *for* statisticians and not exclusively at training *of* statistics.

2.2 *Targets groups*

The programmes are designed for public sector statisticians in the Member States of the European Community and EFTA. However, the programmes will also be open for the National Statistical Institutes of countries in Central and Eastern Europe and private sector statisticians. The courses and the seminars thus have a truly international audience.

2.3 *Training philosophy*

The training programmes in the TES project are vocational training programmes. This means that the training modules are directed at professional categories of statisticians and not at university students. Training aimed at professional categories can rely much more on real-world examples than university training can. In fact it should take the problems encountered by the participant in his working environment as point of departure. This approach allows to attract a direct interest of the participant who will realise the immediate utility of a particular training module for his professional work. Hence, the pedagogical method is problem-oriented and based on the principle of learning-by-doing whenever possible.

2.4 *Programme and curriculum development*

The content of the TES training programmes can be defined at three different levels of detail. At a global level one may look at the composition of the programme in terms of the various main areas of training (e.g. statistics, automation, etc.). Secondly, one may consider the choice of training modules within each main area. Lastly, the content of each training module may be examined. For each of these levels of detail the leading principle is that choices are made on the basis of the training needs in the countries concerned.

Every year, the proposals for the annual training programme are submitted to the Scientific Committee (SCICOM) of the TES project. The main task of the SCICOM is to maintain the relevance of the training programme in view of the training needs perceived.

Moreover, the TES-Secretariat regularly conducts a survey of training needs of the National Statistical Institutes of the Member States of the European Community and EFTA. The TES Secretariat also identifies the areas in which National Statistical Institutes and academic institutions can contribute to the training programme.

The TES programmes cover the following main training areas:

I Statistical Techniques (TEC)
II Official Statistics (OFF)
III Economic and Social Analysis (ESO)
IV Automation (EDP)
V Management and Policy (MAP)
VI European Integration (EUR)

The training programmes distinguish between two types of modules: *courses* which are modules aimed at junior and mid-career statisticians with a duration varying between 3 days and 3 weeks and *seminars* which are modules aimed at senior statisticians and managers with a duration of 2-3 days. Courses may be regular (i.e. reappear in the subsequent yearly programmes) or may have an ad-hoc character. Seminars are modules in which the preponderant learning mode will be the exchange of experience and know-how on the basis of contributed papers. Seminars may either have a scientific character, bringing together specialists in a specific field, or they may be policy - or management - oriented, bringing together top executives and managers of Statistical Institutes.

The annual training programmes contain a menu of courses and seminars from which one may select freely, provided the entry qualifications are satisfactory. The courses are given in English, French or German or are multilingual (simultaneous interpretation with two or three of the before mentioned languages). The seminars are all multilingual.

The policy with respect to the curriculum development in each main area is based on the following criteria. For the annual programmes the choice of the training modules is based on:

a) the expressed needs of the National Statistical Institutes;
b) the number of registrations in the previous programme;
c) the results of ex-post evaluation by participants and teaching staff of modules taught in previous programmes;

d) assessment by the sponsoring organisations of the effect of training modules on the professional performance of their staff;
e) the results of discussion in the SCICOM (Scientific Committee of TES).

Furthermore, the contents of each course, both in terms of subject matter and in terms of training method are subject to regular revisions based on participants' assessment, evaluation of teaching staff and assessment by the NSIs.

The continuous-feedback of the sponsoring institutions, participants, teaching staff and the SCICOM is considered essential for the programme to remain relevant.

3. The experience of the first three years

3.1 *The programmes*

The overview of the courses together with their frequency, their duration and the language in which they were given is presented in Table 1 of the annex. Detailed contents for each of the courses can be obtained from the TES Secretariat. The courses and the seminars mentioned in the annual programmes are given in different locations all through Europe.

3.2 *The participation*

Until now the participants in the annual TES training programmes have been public sector statisticians from the all Member States of the EC and the EFTA and from the following transition countries (referred to as ECO): Bulgaria, Hungary, Poland, Romania, Slovak Republic, Czech Republic and Slovenia (Yugoslavia participated only in 1991).

The total number of persons participating in the courses of the annual programmes of 1991-1992, 1992-1993 and 1993-1994 are 275, 367 and 532 respectively. This number does not contain the participation in seminars which varied strongly over the different programmes due to the variation in number of seminars programmed.

The tables 2 to 4 of the annex provide for each of the three annual programmes details on the participation per region and per main training area. Fig. 1 shows that when the training programme started in 1991 more than half of the participants (58%) attended courses related to European Integration issues (EUR); the second important training area was Statistical Techniques, followed by Automation (EDP) and Official Statistics (OFF) which, for this programme, was restricted to national accounts statistics. The distribution of participants between the various training areas is obviously related to the number of courses offered in each training area.

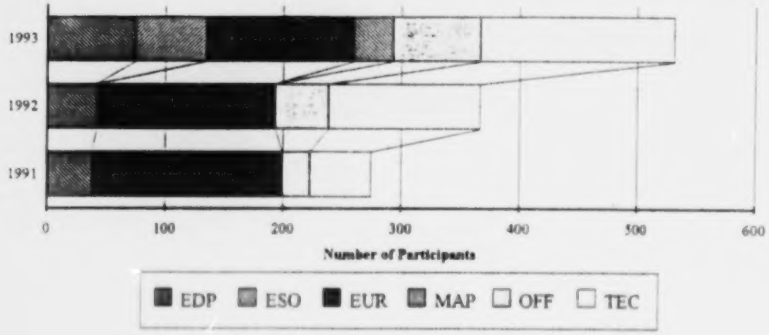


Figure 1. *Number of course partecipants per main training area*

However, it is also related to the average class size in each of the main areas. For the more technical training areas the average class size has to be significantly lower than for the more general subjects. Fig. 2 gives an impression of these differences.

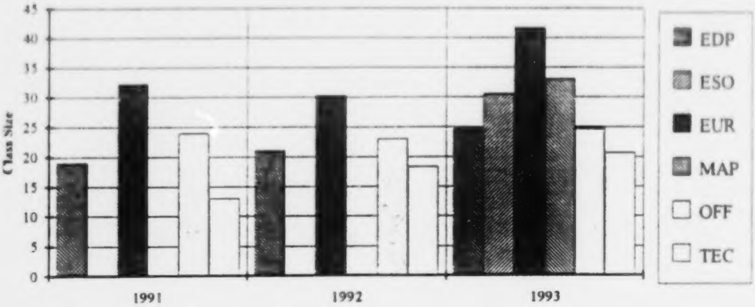


Figure 2. *Average class size per main training area*

In the programme 1992-1993 the number of courses on Statistical Techniques was increased from 4 to 7, implying that the participation in this type of courses increased from 19 to 35 per cent. Fig. 1 also shows an increase for Official Statistics related to the introduction of second course in French on national accounts statistics.

Finally, the programme 1993-1994 was restructured. It includes two new training areas, Management and Policy (MAP) and Economic and Social Analysis (ESO), bringing the number of main training areas to five. Moreover, this programme only counts two courses on the Functioning of the EC and its Statistical System instead of three in the previous programmes. Consequently, the distribution of participants over the training areas changes significantly, leaving Statistical Techniques the largest training area with 165 participants (32%), followed by European Integration with 125 participants (23%) and Automation and Economic and Social Analysis with both 74 participants (14%).

Also interesting is the distribution of participants according to their region of origin. Fig. 3 shows that the training project started with a very high proportion of participants coming from EC countries and relative small proportions coming from EFTA and Central European (ECO) countries. The implications of the political changes in Europe around the beginning of the decade on the TES training programme are evident: by 1993 almost half of the participants came from ECO countries, slightly more than half from EC countries and about ten per cent from EFTA countries. The overall growth in participation was about 25% per annum, that of EC participation 4%, that of EFTA participation 63% and that of ECO participation 83% per annum.

The participation in seminars was very high in the first year of operation as there were four seminars organised that year. For the subsequent programmes only one seminar per year was programmed.

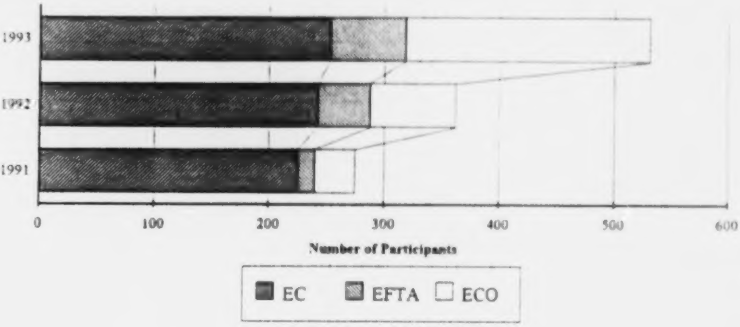


Figure 3. *Number of course participants per region*

3.3 *The training method*

The choice of training methods in this professional training programme for statisticians is a factor of paramount importance for the success of a course. It is not always easy to convince the training staff that traditional lecturing as a one-way transfer of know-how is not the best way in adult training.

Due to the fact that most statistical institutes are not prepared to let their staff attend courses for a long period most courses have a duration of about 5 days. This is a very short time span if one wants to give *theory* the place and the time it needs. Therefore the present trend is to provide the future participants with the necessary theoretical course material well in advance and to require a sound theoretical preparation before attendance of the course. In this way the course staff can concentrate on the discussion of real world case studies (often based on empirical material provided by the participants) and on *hands-on* training using increasingly personal computers and standard statistical software packages.

Practical work during the courses is often performed in small subgroups and requires a high trainer-participant ratio.

A special seminar for trainers in the TES programme on modern adult training methods is planned to take place at the end of 1993.

4. Conclusions and perspectives

In conclusion it can be said that this initiative of Eurostat has shown to cater for a real need for training and to provide a good platform for international contacts between European statisticians.

For the future development of the project "Training of European Statisticians" the following orientations and initiatives will be given priority:

- Continuing attention will be paid to the practical applications and learning-by-doing in the various courses. The applications should, as much as possible, relate to the day-to-day work in the national statistical offices both at national level and in their relation to Eurostat.

- Case studies will be edited both at national and Eurostat level which could serve as illustrations and data bases for applications in various courses. These case studies should have a project character and cover a range of aspects specific to a statistical project.

- The TES Secretariat will continue developing course manuals that have a direct utility for the participants in their work environment.

- The TES programme will continue to play a supportive role with regard to the harmonisation of methods and concepts in the EC.

- In principle, participants from countries outside Europe can be given the opportunity to participate in TES courses provided that their candidature is supported by a NSI of an EC or EFTA Member State. Applications will be considered case by case and competition with other training centres (CESD network) will be avoided.
- Measures will be taken to improve the homogeneity as regards the background of participants attending the same course. Entry requirements will be made more specific. The correspondents will be asked to do their utmost to stress the importance of the fulfilment of the entry requirements both with the participants and the authorities responsible for the selection of the participants of their institutions.
- It has been observed that the balance between theory and applications is not a sufficient condition for the courses to be fully appreciated by the participants. The training methodology should be oriented towards the needs of professional adult training, i.e. the introduction of theory and methodology should be carried out in the context of a problem oriented exercise. Several training techniques with this objective have been developed and it was decided to organise a *seminar for TES trainers* in the second half of 1993. The objectives of the proposed seminar are: to offer a platform of discussion and exchange of experiences concerning training methods applied so far in the annual TES training programmes, (participants are invited to submit short position papers), to provide the trainers with expert views on practical solutions for training methods specifically aimed at professional training on the basis of a number of invited papers, to offer the trainers the opportunity to discuss the problems encountered in specific learning situations with experts in training method. The seminar should result in practical guidelines on training methods for TES trainers.
- With respect to target groups the prime orientation will remain towards the NSIs. However low-cost promotion will be made towards networks of various kinds, e.g. chambers of commerce. The publicity of the TES programme should also be extended to regional and local statistical agencies.

The major challenges for the future are to ensure a high professional standard of the training, to maintain the relevance of the programme, i.e. to follow closely the training needs of the National Statistical Institutes and to continue efforts to better adjust training methods to the needs of our target group.

ANNEX: SOME STATISTICS ON THE THREE ANNUAL PROGRAMMES AND INFORMATION ON THE 1994-1995 PROGRAMME

Table 1. *Overview of the three programmes and the planned 1994-1995 programme*

COURSE/SEMINAR TITLES	1991-1992			1992-1993			1993-1994			Planned 1994-1995		
	Course Frequency	Language(s)	Duration (Days)	Course Frequency	Language(s)	Duration (Days)	Course Frequency	Language(s)	Duration (Days)	Course Frequency	Language(s)	Duration (Days)
Automation (EDP)												
Application of the BLAISE system	1	E	5	1	E	5	1	E	5	1	E	5
Relational Databases							1	E	3	1	E	3
Application of the SAS System	1	E	5	1	F	5						
Economic and Social Analysis (ESO)												
Demographic Data Analysis							1	E	5	1	E	5
The Construction and Use of Level of Living Indicators							1	E	5	1	E	5
European Integration (EUR)												
The Functioning of the EC and its Statistical System	3	M	5	3	M	3-5	2	M	3	2	M	3
The European System of Integrated Economic Accounts (ESA)	1	M	5	1	M	5						
The Implementation of the Revised European System of Integrated Economic Accounts (ESA)										2	EFD	5
The INTRASTAT System				1	M	3	1	M	3	1	M	3
Using Eurostat Statistics							1	F	3	1	F	3
Pilot Surveys for the Service Sector										1	M	3
Management and Policy (MAP)												
Principles and Techniques of Statistical Project Management							1	E	4	1	E	4
Official Statistics (OFF)												
National Account Statistics in Practice	1	E	8	2	EF	8	1	E	8	1	E	8
The Set-up and Use of Registers							2	DE	5	2	DE	5
Statistical Techniques (TEC)												
Exploratory Data Analysis	2	FD	5									
Sampling Techniques and Practice	1	F	15	2	EF	15	2	EF	15	2	EF	15
Applied Time Series Analysis	1	E	5	1	F	5	2	EF	5	2	EF	5
Introduction to Applied Time Series Analysis										1	E	5
Advanced Methods in Applied Time Series Analysis										1	F	5
Index Number Theory and Practice				1	D	5	2	EF	5	1	F	5
Theory and Application of Household Panel Surveys				1	E	9	1	E	9	1	E	9
Measurement in Surveys				1	E	5	1	E	5	1	E	5
Dealing with Non-Response							1	E	3	1	E	3
Classifications: One Language and one Structuring							1	F	5	1	F	5
Seminars												
International Comparison of Survey Methodologies	1	M	3									
Statistical Project Management	1	M	3									
Strategic Issues in Statistical Policy	1	M	3									
Dissemination Policy for NSI Outputs	1	M	3									
Quality Control of Statistics				1	M	3				1	M	3
Environmental Statistics							1	M	3			

Note: D = German, E = English, F = French, M = Multilingual (as a rule English, French and German)

BEST COPY AVAILABLE

Table 2. *Summary of participation in 1991-1992 programme*

	Number of Courses	PARTICIPATION PER REGION AND PER MAIN TRAINING AREA											
		TOTAL			EC			EFTA			ECO		
Number of Countries		25			12			5			6		
Training Areas		Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country
EDP	2	38	19.0	1.5	32	16.0	2.7				6	3.0	1.0
ESO													
EUR	5	161	32.2	6.4	136	27.2	11.3	10	2.0	2.0	15	3.0	2.5
MAP													
OFF	1	24	24.0	1.0	11	11.0	0.9	4	4.0	0.8	9	9.0	1.5
TEC	4	52	13.0	2.1	46	11.5	3.8	1	0.3	0.2	5	1.3	0.8
TOTAL	12	275	22,9	11.0	225	18.8	18.8	15	1.3	3.0	35	2.9	5.8
SEMINAR	4	152	38.0	6.1	119	29.8	9.9	15	3.8	3.0	15	3.8	2.5
GRAND TOTAL	16	427	26.7	17.1	344	21.5	28.7	30	1.9	6.0	50	3.1	8.3

Table 3. *Summary of participation in 1992-1993 programme*

	Number of Courses	PARTICIPATION PER REGION AND PER MAIN TRAINING AREA											
		TOTAL			EC			EFTA			ECO		
Number of Countries		25			12			7			6		
Training Areas		Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country
EDP	2	42	21.0	1.7	27	13,5	2,3	5	2,5	0,7	8	4.0	1,3
ESO													
EUR	5	151	30,2	6.0	108	21.6	9.0	22	4,4	3.1	18	3.6	3,0
MAP													
OFF	2	46	23,0	1,8	33	16,5	2,8	4	2.0	0.6	9	4,5	1.5
TEC	7	128	18.3	5.1	74	10,6	6,2	15	2,1	2.1	39	5,6	6.5
TOTAL	16	367	22.9	14,7	242	16,5	20.2	46	2.3	6.6	74	11.3	12.3
SEMINAR	1	55	55,0	2,2	34	34.0	2,8	10	10.0	1,4	7	7,0	1,2
GRAND TOTAL	17	422	24,8	16,9	276	16,2	23,0	56	3,3	8,0	81	4,8	13.5

BEST COPY AVAILABLE

Table 4. *Summary of participation in 1993-1994 programme*
(preliminary figures)

	Number of Courses	PARTICIPATION PER REGION AND PER MAIN TRAINING AREA											
		TOTAL			EC			EFTA			ECO		
Number of Countries		25			12			6			7		
Training Areas		Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country	Total Nr	Per Course	Per Country
EDP	3	74	24.7	3.0	27	9.0	2.3	6	2.0	1.2	41	13.7	6.8
ESO	2	61	30.5	2.4	21	10.5	1.8	7	3.5	1.4	33	16.5	5.5
EUR	3	125	41.7	5.0	87	29.0	7.3	11	3.7	2.2	27	9.0	4.5
MAP	1	33	33.0	1.3	16	16.0	1.3	1	1.0	0.2	16	16.0	2.7
OFF	3	74	24.7	3.0	33	11.0	2.8	10	3.3	2.0	31	10.3	5.2
TEC	8	165	20.6	6.6	69	8.6	5.8	30	3.8	6.0	66	8.3	11.0
TOTAL	20	532	26.6	21.3	253	12.7	21.1	65	3.3	13.0	214	10.7	35.7
SEMINAR	1	16	16.0	0.6	5	5.0	0.4	2	2.0	0.4	9	9.0	1.5
GRAND TOTAL	21	548	26.1	21.9	258	12.3	21.5	67	3.2	13.4	223	10.6	37.2

Note
EDP = Automation
ESO = Economic and Social Analysis
EUR = European Integration
MAP = Management and Policy
OFF = Official Statistics
TEC = Statistical Techniques

EC = Member States of the European Community
EFTA = Member States of the European Free Trade Association
ECO = Transition Countries Central Europe

BEST COPY AVAILABLE

TEACHING BIOSTATISTICAL CONCEPTS AND METHODOLOGY TO MEDICAL PROFESSIONALS: VARIOUS ALTERNATIVES

Shrikant I. Bangdiwala
Department of Biostatistics, CB#7400
University of North Carolina, Chapel Hill, NC 27500-7400, USA

1. Introduction

Physicians and other medical professionals are increasingly aware of their need for biostatistical knowledge, not only if directly involved in research activities, but also if, as a clinical practitioner, one wishes to keep abreast of advances in the field.

The traditional approach involves incorporating biostatistical research concepts within the medical school instruction. The curricula in most medical schools in the United States and in other countries now include at least a session on introduction to biostatistics, although it may be called epidemiological research or research methods. Such courses are typically short (one rotation or one semester) and cover basic information with the intent of introducing the general concepts of descriptive and inferential statistics, but are not intended to prepare the physician for conducting clinical research. Most medical students do not envision a career as investigators but as practitioners, and thus tend to concentrate on the traditional medicine subjects and therefore to disregard the importance of the statistical instruction. It would seem that the timing of this instruction should come at a stage in their medical career when the professional accepts the importance of basic biostatistical knowledge.

For the few medical professionals who do choose a career in research, it is imperative that additional in-depth research methods instruction be obtained. In some cases, they are able to pursue a post-graduate master's of public health (MPH) degree, typically in a department of epidemiology. The University of North Carolina at Chapel Hill offers MPH degrees in the Departments of Biostatistics and of Epidemiology of the School of Public Health. However, many physicians do not have the opportunity to pursue such post-graduate training and may either be self-taught or often disregard the need for further training and merely "forge-ahead". This latter approach has been greatly facilitated since the last decade with the widespread availability of statistical software for personal computers, so that anyone "armed" with these tools can be their own biostatistician. Numerous examples exist of the pitfalls of such activities.

Another alternative is to rely on a biostatistician and to blindly accept the answers obtained from such collaboration. This is not a healthy approach either as the biostatistician is placed on a pedestal, viewed as a necessary evil that must provide significant results, and there is no semblance of a collegial collaboration.

The end results of such alternatives is often poorly planned, poorly conducted, poorly analyzed or poorly presented medical research. It is in the best interest of the medical research profession and the biostatistical profession that medical professionals planning to be clinical researchers obtain relevant, targeted statistical education. This is not to argue that the biostatistically-trained medical professional should be able to be their own biostatistician, but that if they have a basic understanding of the statistical issues present in their research, they will be able to better collaborate with a statistician and, therefore, to be better researchers.

This paper presents four non-degree post-graduate alternative statistical education approaches that aim to provide medical professionals with the biostatistical research tools necessary to be effective clinical investigators. They range in coverage, time commitment, cost and depth of training. The goals of each approach are presented and compared, especially in terms of the expectations of the biostatistical abilities of the course participants upon completion of training. The experience of the author as instructor in all four different teaching approaches is presented.

2. Clinical Scholars training: two years in-depth

The Clinical Scholars training program is a Robert Wood Johnson Foundation funded program in the United States that competitively selects physicians nationwide for a two-year post-graduate training program in clinical research methods. Fellows attend training centers within selected medical schools at research universities. This training is quite similar in time commitment to pursuing an MPH degree, except that it is offered by a school of medicine and is more focused on clinical research rather than epidemiology. Its goal is to produce excellent clinical researchers.

The biostatistical training provided at the University of North Carolina through this approach is divided into two components. Fellows take a one-year intensive core curriculum of research methods, in which ten sessions are directly on basic biostatistics. These sessions compress a one-semester introduction to probability, descriptive and inferential statistics into 30 hours of instruction. In addition, the remainder of the core curriculum covers other biostatistical topics indirectly as they pertain to specific study designs like case-control studies or clinical trials. The second component comes from enrollment in regular graduate level courses in the Departments of Biostatistics or of Epidemiology, where fellows take intermediate subjects like multiple regression, clinical trial methodology and multivariate methods.

This approach is quite costly in terms of time commitment on the part of the physician, but has demonstrated success in producing excellent clinical researchers. They are still dependent on biostatistical support, but are conversant with statistical terminology and concepts. They are able to critically appraise the medical literature.

3. International Clinical Epidemiology Network training: one year in-depth

The International Clinical Epidemiology Network (INCLEN) training program is originally a Rockefeller Foundation funded worldwide program that competitively selects physicians from 26 selected medical schools in developing countries for a one-year post-graduate training program in clinical research methods (Halstead *et al.*, 1991). Fellows attend training centers within selected medical schools at research universities in Canada, Australia or the United States. This training is similar in intent to the Clinical Scholars program, but with a reduced time commitment.

The goals of the INCLEN program are to develop units of excellence in clinical epidemiology research at the participating medical schools in the developing countries. As such, the program not only provides training to the future investigators, but also time protection for conducting research activities and the necessary biostatistical and other support required upon their return to their home institutions.

The biostatistical training provided at the University of North Carolina through this approach is almost identical to the Clinical Scholars program. The reduced time-in-residence limits the possibility of more in-depth biostatistical courses that the physicians can take. The medical professional thus may not be as independent statistically as an MPH in epidemiology or as a Clinical Scholar, but the Network is designed so that the necessary biostatistical support is available at the home Clinical Epidemiology Unit (CEU).

This approach is slightly less costly in terms of time commitment, and fellows are prepared to not only be critical appraisers of the medical literature, but to function as independent clinical researchers within a team at the CEU.

4. Clinical Vision Research short course: two-day introduction

The National Eye Institute of the United States' National Institutes of Health sponsors two-day intensive short courses in clinical vision research methodology for ophthalmologists in developing countries. These short courses are offered in conjunction with major regional scientific professional meetings so as to attract a broad-based audience.

The goals of this course are to introduce ophthalmologists to clinical research methods and develop skills to evaluate clinical research protocols. This is achieved through an overview of basic epidemiologic and statistical principles, and seminars/discussions on the major issues to consider in the design, conduct, analysis and interpretation of clinical research. Participants thus develop a critical approach to clinical research.

The biostatistical methodological topics and concepts are not directly offered in such short courses given the time limitations, but through applications to particular study designs and the conduct and analysis of epidemiologic studies. An integral component of the short course is a detailed critical appraisal of two selected ophthalmology research articles.

This approach does have the potential of reaching a large section of the medical professionals in a given field with relevant specific instruction. It requires minimal time commitment and involves less costs as well. However, it also is limited in the depth of statistical understanding that the participants acquire. Participants cannot be expected to gain in-depth appreciation of statistical issues, but should acquire a critical appreciation of medical research in their field. Those participants who wish to pursue careers as researchers would seek further in-depth training.

5. Extended short course: one-month intermediate

The final alternative is an intermediate approach to the ones presented above, an intensive but short course. In the United States, several institutions offer summer continuing education programs with the intent of providing an intensive introduction to epidemiology and biostatistical methods for health professionals planning careers in research, teaching or practice in a particular field. The Council on Epidemiology and

Prevention of the American Heart Association offers a 10-day seminar focused on methods for cardiovascular diseases (19th session in 1993); the New England Epidemiology Institute and Tufts University offers a 3-week series of 14 courses on Epidemiology (13th session in 1993); the University of Michigan School of Public Health offers 1-week and 3-week courses as part of their Graduate Summer Session (28th session in 1993). These are but a few examples of alternative instructional possibilities for acquiring research tools. These relatively short courses are viable alternatives for the medical professional seeking such training but unable to commit larger amounts of time. They are intensive, vary in-depth, and provide continuing education credits as well as graduate level credits if desired. The biostatistical concepts are presented either as a basic introduction course (ex. Fundamentals of Biostatistics), indirectly as part of an applied epidemiology topic (ex. Clinical Trials - Analytic Methods), or as a special in-depth statistical topic (ex. Introduction to the Logistic Model).

The Department of Anesthesia and Intensive Care of the John Hunter Hospital in Newcastle, Australia, chose the intensive 1-month approach for training its clinical residents in the fundamentals of biostatistics in June 1991. They contacted the Centre for Clinical Epidemiology and Biostatistics of the University of Newcastle to offer the course. Four in-class sessions over a 4-week period were conducted, coupled with critical appraisal readings and exercises, with the goal of providing an understanding of commonly used statistical methods encountered in the current medical literature (with an emphasis on anesthesiology). The extension over a 4-week period enabled participants to complete readings and exercises, but also to develop in-depth questions and to better "digest" the statistical topics. It also allowed for minimal disruption of their clinical duties, and by being offered on-site, accommodated potential clinical emergencies.

At the end of the course, participants had a basic understanding of a few commonly used statistical methods, and were better equipped for critically appraising the anesthesiology literature.

6. Discussion and conclusion

The four alternatives presented above all aim to provide the medical or health professional with training in biostatistical methods and concepts. Choosing a particular approach depends on the time commitment on the part of the potential participant, which should be directly related to their interest in conducting clinical research. Those medical professionals wishing to pursue careers in research should obviously choose thorough training, either with a post-graduate degree or competitively applying for training fellowships such as the Clinical Scholars or the INCLEN programs. However, the most practical approach is to participate in short courses, since selection only depends on one's ability to pay (or obtain a scholarship). A 2-day course is limited in its training potential, and should be considered as one to reach a large group of medical professionals with a basic critical appreciation for the importance of biostatistics, but only to stimulate the potential researcher to pursue further training. An extended but still short approach seems the preferred alternative. It is quite flexible and accommodating to the medical professional's schedule, and the statistical content can be varied from the basic introduction, to intermediate topics or even advanced topics. The typical short course summer sessions in the United States offer several courses to choose from. Tailor-made courses for specific medical professions can focus on the statistical issues particular to that field and may be also a valid alternative.

The goal of educating the medical professional consumer of statistics is a lofty one, but attainable. The approaches presented in which the author has been an instructor have all met their stipulated training goals. Training health and medical professionals in statistical methodology is not a threat to the statistical profession, but actually enhances our profession. Given the potential abuses of statistical methods made easy by the proliferation of statistical software readily available, the potential danger to the profession from the ill-trained casual user is great. Thus, an educated consumer is the best client for our statistical profession.

Bibliography

Halstead S.B., Tugwell P. and Bennett K. (1991), The International Clinical Epidemiology Network (INCLEN), A progress report, *Journal of Clinical Epidemiology*, 44(6), pp. 579-589.

TEACHING STATISTICS TO PRODUCTION PERSONNEL

Andrés Carrión Garcia
Department of Statistics, Universidad Politecnica de Valencia
Cno. de Vera s/n., 46020 Valencia, Spain

1. Introduction

The highly competitive situation existing for many products and in many markets, together with customers' increasing demands (quality, reliability, safety, etc.), has moved companies to improve, among other aspects, their production quality.

In the world of industrial quality, statistics plays an essential role. As Deming (1982) says, no other resource is in such short supply in companies as statistical knowledge, and no other can help so much to increase quality and productivity. Guided by this principle many enterprises, following the big international car manufacturing companies have launched training programs for their employees in which, and under different names (Statistical Process Control, Statistical Methods, Statistics for Non-production Departments, ...) statistics-related knowledge is taught. The courses offered vary as widely as the needs that may appear in the industrial environment, from basic statistical knowledge to design of experiments and other advanced techniques. It is also true that, sometimes, methods which are not strictly statistical are included together with statistics, as happens with some of the seven Japanese tools.

2. Contents of the training programs

As suggested before, a wide variety of subjects is included in the different courses. Usually these subjects are grouped under the following headings:

Descriptive methods. Different kinds of tables and graphics are presented, and their adequacy for each situation, construction and interpretation are discussed.

Probability. It is, perhaps, the subject least studied in these courses, except in the advanced ones, where the definition of probability and its basic properties and theorems are discussed. At the other levels, only an intuitive approach is usually presented.

Probability distributions. Among the discrete distributions, special emphasis is given to the Binomial and Poisson because of their application to quality control where variables are classified by discrete attributes. The Normal distribution (very important for quality control of continuous variables) is also studied and, depending on the level of the course, t, F and chi-square distibutions are presented. One other distibution used, but not theoretically analyzed, is the Pareto distribution, which occurs in one of the Seven Tools (the "Pareto Analysis").

Regression and correlation. The concepts of correlation coefficient and simple regression are explained, as well as the idea of the least squares method. Computation of the model parameters and a graphical non parametric method are also touched on.

Comparative studies. Usually centered on the paired test for means and variances (t and F tests). In advanced courses ANOVA techniques are also included.

Design of experiments (DOE) and Taguchi techniques. Starting from the basic concept of the need for initial experimental design as exposed in basic courses, at advanced levels there are monographic courses devoted to DOE, from both classic and Taguchi approaches.

Statistical Process Control (SPC). SPC techniques are a set of estimation and hypothesis-testing methods, combined with a graphical analysis method. The different kinds of control charts, adapted to different process needs, make use of the Binomial, Poisson and Normal distributions. In addition to the standard Shewhart method, and depending on the type of industry, CUSUM and EWMA charts are also explained.

Acceptance control. This course includes techniques for acceptance sampling of the finished product. Usually it is oriented towards the use of standards like MIL-STD 105D or MIL-STD 414 (or their equivalents in Spanish UNE standard) and the Dodge-Romig Tables.

Sampling. Concepts like randomness and representativity of the sample are defined and commented on. Computation of sample size for various circumstances is also studied.

3. Conditioning factors in the students

The main conditioning factor in the persons attending these training programs is their heterogeneity. In concrete terms:

- In the same kind of course, and frequently in the same group, the levels of the students are very different, not only in their mathematical and statistical backgrounds, but also in other general aspects.
- Secondly, and also very often, interest in the subject of the course varies from those who are enthusiasts for a technique they consider as very useful for their job, to those who are forced to attend the course by the management and/or don't find the subject of the course applicable to their present job, and show a lack of interest in or even reject the material.
- Thirdly, there are students who come from different departments of the company, having different approaches and points of view, thereby making it difficult to find examples of general interest.

Another characteristic of the students is what we can call a "strong wish for early application" of the knowledge they receive in the courses. This defines three groups of persons:

- those who find the tools presented in the course of immediate utility and are highly motivated by this usefulness;
- those who find the tools of low applicability to their job and who show lack of interest in and perhaps reject the course;
- and finally, those who, perhaps without seeing any immediate application of the technique to their job, are interested in knowledge and tools that they consider useful.

The attitudes derived from these three groups seriously affect the development of the courses.

Together with these attitudes of the students, and probably a partial cause of some of them, is the attitude of the management, who frequently demand immediately applicable results from the courses, by means of suggestions, improvements, studies, etc., sometimes confusing training with consulting.

4. Methodological outlines

The characteristics of the students and of the matter to be taught, plus the fact that this is non-academic teaching, have determined a certain approach and teacher attitudes adapted to such factors. The main characteristics are:

- Reinforcement of the intuitive approach rather than the rigorous mathematical approach, both in the definition of concepts and in the presentation of techniques and analysis of results.
- Frequent use of examples taken from the immediate environment of the students (to increase the interest in and motivation by the course).
- Importance of individual and team work of students using real data obtained by themselves.
- Use, limited by the heterogeneity of students, of different computation devices (pocket calculator, computer) to relieve students of tedious repetitive calculations and make the use of more advanced, and complex techniques possible. Related with this point, it is interesting to mention the difficulties we find in using informatics, due to the fact that, especially in some levels, students have very little experience in this field, requiring prior introduction sessions. Only high level courses are free of this problem.
- The need to have clear and specially adapted material, for both examples and the content of the course.
- Use of simple language, avoiding technical terms and complicated constructions.
- When mathematical expressions are used, placing more emphasis on the sense than on the expression itself: explaining the formula.

5. Positive and negative aspects of the training programs for production personnel

We will speak briefly about advantages and disadvantages that training activities like these have for companies, the personnel receiving the courses and the university professors involved in them.

For the companies, collaboration with universities facilitates access to specialists in areas usually not found in the industry. In the area of statistics this is especially true (remember the sentence by Deming quoted before). Sometimes it can be difficult to gain acceptance of the contents and approach to the courses from university professors, who have the habit of making wide use of a mathematical approach to statistics.

Personnel attending the courses have a very good opportunity of receiving complementary training, which is sometimes very limited, in areas of interest to their job. They receive information on new methods and revise their way of working, adapting it to technical developments and to company policies. Inconveniences may result from poor selection of the persons attending each course, causing difficulties to the whole group.

For the university professors involved, these training programs are good opportunity to demonstrate the practical uses of their knowledge, to apply this knowledge themselves and to adapt their interests (related to teaching and research) to society's demands. Our university produces mainly engineers in different branches, and thus it is very important to keep teaching in contact with reality. There is a real danger of using a level (mathematical, technical, ...) of the industry training programs in the university courses, that could cause, from our point of view, a loss of quality in academic teaching. In any case, the advantages are greater than the disadvantages. University teaching is improved by different factors: firstly, the need for conceptualizing, frequent in industry training, can be, to some extent, used in some academic courses. Secondly, the contents of university courses change to fit industry needs, improving the future engineers' background. Thirdly, there will be many real examples of the use of statistical methods, increasing the motivational level of students (Barnett, 1983). And fourthly, real experience in industry causes the appearance of new research fields that can lead to new developments, as has happened frequently in the history of statistics and operations research.

Bibliography

Barnett V. (1983), *Statistical Education*, Dirección de Estadística del Gobierno Vasco, Vitoria, Spain.

Deming W.E. (1982), *Quality, Productivity and Competitive Position*, Cambridge University Press, Cambridge.

ISEC, CALCUTTA:
THE PIONER CENTRE FOR STATISTICAL TRAINING OF GOVERNMENT SPONSORED STATISTICIANS FROM DEVELOPING COUNTRIES OF THE THIRD WORLD, MAINLY FROM ASIA AND THE FAR EAST

A.B. Raha
International Statistical Education Centre
203 Barrackpore Trunk Road, Calcutta 700035, India

1. Introdution

In 1916 W.T. King in his book (King, 1916) attributes the appellation "Father of Statistics" to Gottfrid Achenwall of Germany who first applied the term "statistics" in 1746 to a series of lectures primarily about data concerning the land, its products, the Government, state, resources and the like. But it was Karl Pearson and R.A. Fisher who laid down the theoretical foundations of modern statistics which enabled statistics to emerge as a new discipline divorced of its original pure descriptive background. While this is the global scenario of modern statistical development, the name that comes up first in the context of statistical education in Asia is that of P. C. Mahalanobis. The happy coincidence of the birth centenary of P.C. Mahalanobis with the first scientific Meeting of the newly formed IASE provides a suitable forum to assess his contribution towards dissemination of statistical knowledge in Asia. Mahalanobis realised the importance of statistics in the twenties during the nascent stage of its development in the West and recognised statistics as the "key technology" in planning for the economic development of independent India. He pioneered the statistical movement in India, in fact in Asia at large, by founding the Indian Statistical Institute in 1931, the National Sample Survey (NSS) and the International Statistical Education Centre (ISEC) in 1950. He truly guided India to assume the leadership in the spreading and development of statistical education among Asian countries. Last but not least, the influence of Fisher on Mahalanobis and the interdependence of these two statistical giants of this century have the greatest impact on the statistical system in India and, *ipso facto*, on the statistical education of Asia as a whole.

2. Birth of ISEC and SIAP

After this short prelude, I come directly to the aspect of statistical education in Asia I plan to dwell upon in this write-up. I would like to focus on the training of staff for government statistical offices. The views expressed are formed by my experience of running the various activities of ISEC in the capacity of its Member-Secretary, Board of Directors over last six years. It may be recalled that after the establishment of the International Statistical Institute (ISI) in 1885 its small initial membership consisted mainly of official statisticians, and an important activity of ISI in the past 107 years has been the upgrading of statistical procedures and organisations, particularly in the third world countries, by support for the training of government statisticians in international statistical training centres. Among these the earliest and the most important one in Asia is the International Statistical Education Centre (ISEC) at Calcutta. P.C. Mahalanobis felt after Indian independence the need to bridge the gap between academic and official or government statistics for socio-economic development of the country. This urge led him to set up ISEC at Calcutta within the precincts of the internationally famous Indian Statistical Institute in 1950, which was recognised and is being operated by ISI jointly with the Indian Statistical Institute, under the auspices of UNESCO and the Government of India. Since inception the centre has been providing statistical training to government sponsored statistical staff from countries in the Middle East, South and South East Asia and the Far East. Its coverage was later expanded to the Commonwealth countries of Africa. At present ISEC is the only centre of its kind which is directly under the supervision and control of ISI.

Among other international statistical training centres that attract international participation is the Statistical Institute for Asia and the Pacific (SIAP) at Tokyo. A brief account of the genesis of this institution which has been in existence for the last two decades may be of some interest. Wide appreciation of the need for a regional training institute in the field of statistics was stimulated by the recommendations of the sixth session of the Conference of Asian Statisticians held in December 1964. In order to remedy the critical shortage of trained personnel at the professional level in many Asian countries, the Conference urged the establishment of a statistical training centre for the region so as to promote training conducive to development of statistics and to bring about regional cooperation in statistical activities. In March 1965, the 21st session of the ECAFE (now ESCAP) endorsed this recommendation with the formation of an Expert Group, with C.R. Rao as chairman, to advise on the concrete steps to be taken in this direction. In January 1969, the UNDP approved the proposal of the ECAFE to establish the Asian Statistical Institute (ASI) at Tokyo as an autonomous inter-governmental institution under the aegis of ECAFE. The Institute was formally inaugurated in 1970 and assumed its present name SIAP in September 1977.

The SIAP General Course lasts six months and is a thoroughly professional course in statistics. It accepts applications from all countries in the ESCAP region subject only to availability of funding and slots, which are usually 30 a year. English is the medium of instruction in both SIAP and ISEC.

3. ISEC experience

The ISEC offers a 10-month (June to March) Regular Course of training every year leading to a Statistical Training Diploma. The courses are divided into two parts. The first eight months are devoted to training in general statistical methods including training in official statistical systems, the latter being conducted by the Central Statistical Organisation of the Government of India. During the last two months of the 10-month period, each trainee specialises in one selected branch of applied statistics including data processing depending upon the background and need of the trainee arising from the type of Government Organisation he/she is attached to. As mentioned earlier ISEC Calcutta receives applicants from a region wider than that of ESCAP. The average number of participants per year is 25-30, the number being limited by the availability of financial support from various sources. A majority of the trainees have been supported by fellowships awarded by the Government of India, mainly under the Technical Cooperation Scheme of the Colombo Plan, the Special Commonwealth African Assistance Plan, Aid to Sri Lanka and the Indian Technical and Economic Cooperation scheme. The United Nations and its specialised agencies and the Commonwealth Fund for Technical Cooperation, Commonwealth Secretariat, London have awarded fellowships to a few trainees. Since its inception in 1950 the Centre has provided training to about 1200 trainees representing a panorarna of 56 countries mostly belonging to the Third World. Developed countries like Japan, Australia, New Zealand have initially trained their government statistical staff at ISEC and maintain high regard for the training offered at this centre. The countrywise distribution of the trainees is shown below.

Afghanistan	18	Israel	1	Solomon Islands	1
Australia	1	Japan	26	South Africa	1
Bangladesh	22	Jordan	1	Sri Lanka	131
Bhutan	26	Kenya	4	St. Christopher	
Brunei	16	Republic of Korea	41	Nevis-Anguilla	2
Burma	43	Lao PDR	3	Sudan	2
Kampuchea	3	Malaysia	20	St. Lucia	1
Republic of		Malawi	4	Swaziland	4
China (Taiwan)	5	Mauritius	2	Syria	1
Peoples' Republic		Maldives	25	Seychelles	1
of China	4	Nepal	15	Tanzania	36
Ethiopia	6	New Zealand	1	Thailand	53
Fiji Islands	19	Nigeria	55	United Arab	
Gambia	10	Pakistan	64	Republic	2
Ghana	3	Philippines	89	Uganda	13
Hong Kong	15	Papua New Guinea	2	UAE (Abu Dhabi)	4
India	229	Sarawak	1	Vietnam	19
Indonesia	64	Singapore	17	Western Samoa	2
Iran	21	Sierra Leone	9	Zambia	31
Iraq	5	Somalia	1	Zimbabwe	1

The Regular Course is an intensive course with hardly any break in between. The response to the Course is increasing every year as is manifest by the fact that the number of applicants has more than doubled since the ten-month Regular Certificate Course was transformed to a 10-month Diploma Course. The fascination for a diploma or degree among the government statisticians for the betterment of their career acts, quite likely, as an incentive to undergo this course. Moreover, the academic flavour of the course content is an added advantage of the ISEC course over a shorter pure professional course in the sense that the successful trainee feels better equipped to take up higher studies subsequently. In addition to the Regular Course at ISEC, a few government statisticians are admitted on an individual basis, for special courses of varying durations and in different subject fields. Electronic data processing with statistical packages merits special mention.

A number of difficulties are associated with the training programme of the official statisticians, at the Centre. In view of varying needs of countries in the Third World for training statisticians, the ISEC has set only some minimum prerequisites for admission. They must however be able to follow classroom instruction in English and express themselves clearly through spoken and written English. But the language problem is almost always there with such a diverse group of international participants representing usually 15-20 countries each year. Besides the language, the heterogeneity in academic and practical backgrounds of the selected participants is a major deterrent to conducting the training in the most efficacious manner. As a remedy the applicants should preferably be graduates with a few exceptions like Bhutan, Indonesia, Lao PDR, Maldives, Pacific Islands and few others from the Third World where formal education either does not exist beyond secondary school level or is too expensive to be accessible to most of the nationals. As for instance, government statistician participants from Sri Lanka and India are, more often than not, master's degree holders while those from Maldives, Bhutan Pacific Islands and the like have only secondary school background and have to put in hard work to cope with the general standard of the Regular Course. Many developing countries in Asia and Africa or more precisely, in the Third World neither have national facilities to train their official statisticians nor do they have resources enough to send the officials abroad for training without foreign financial assistance. All along ISEC has been ostensibly playing a premier role in imparting training to numerous statistical personnel from these countries. Considering the current trend it is likely that ISEC will continue to enjoy the same status indefinitely as it has been doing since 1950 and maintain the high reputation of Indian statistical education the legacy left behind by P.C. Mahalanobis.

Bibliography

King W.T. (1916), *The elements of statistical method*. The Macmillan Company.

PROBLEMS OF STATISTICAL EDUCATION IN CROATIA

Ante Rozga
Faculty of Economics, University of Split, 58000 Split
Radovanova 13, Croatia

1. Introduction

The process of statistical education in Croatia can be considered to be as complex as it is in all other developing countries. As a part of former Yugoslavia, Croatia was a communist country with an inherited system of secondary and university education. However, the country is also in the process of economic development. Statistical education has been based mostly on the style of the Soviet Union, with the emphasis on statistical theory rather than practice. The available official statistics have been exclusively focused on planning the economy, especially when the new Yugoslav economy was in its infancy. These approaches to statistical education and practice lead to many problems in everyday life. As the Croatian economy is shifting towards a market-oriented economy, statistical research and applications must play a much greater role in all aspects of Croatian everyday life and scientific research. Now, the statistical treatment of scientific problems is spreading everywhere.

2. The lack of practical statistical research in the social sciences

A big problem arises from the fact that there have not been so many surveys in social science as would normally be the case in Western societies. This was due to the anxiety of the ruling party about free interviewing. Surveys about people's political preferences were almost unknown, so people had had no real insights into this until after the first free elections which were held in Croatia three years ago. We conducted one survey among the secondary school students (seventeen and eighteen years old) eight years ago and experienced public consternation when the results showed a high percentage of religious children. The publication of the survey results was later forbidden by party officials as it was not in accordance with the generally accepted opinion that there was a low percentage of religious students.

There have been some surveys about household expenditure, tourist expenditure and other non-political issues, but not about the most important questions of the nation. The statistics for social science has been developing rapidly only in recent years, and there is now an explosion of sample surveys in all fields of social sciences. This sudden explosion, however, has led to misuse of statistics, to many kinds of measurement errors as well as to sundry other errors of methodology and interpretation. Many surveys have been conducted by non-trained people, with little knowledge of statistical theory and methods.

There is also the question of scientific honesty, since some scientists used to "adjust" the results of research according to the expectations of those who ordered the research. It happened especially in social science in a broad sense. So, we need to establish an institute for social research independent of government funding and independent of any political or other party.

For these reasons, we do not have many trained statistical experts in survey sampling, and it will take much effort to produce them in the near future.

For some applications of statistics the process of statistical education is just at the beginning. This is specially true for chemometrics, environmetrics, bibliometrics, scientometrics, but not for computational statistics or econometrics. Statistical research is now expanding in medicine and some other natural sciences, where it was also neglected in the past.

3. Statistics in business and production

There are also major problems in statistical education and research which are connected with business statistics and with the use of statistical methods in improving the quality of production. Since the economy was a so-called "self-governing" socialist economy, the business statistics were also adapted to that kind of production system. The Croatian economy was not well developed, so the use of modern statistical methods was not widespread. Statistical quality control was used only in some big enterprises. In planning the economy, statistics, and especially econometrics, were sometimes treated as something not very necessary. Only in a market-oriented economy can the role of econometrics be totally appreciated.

When measuring the productivity level, prices, salaries and other business variables, there were lots of misuses of statistics reflecting the politics of the day, e.g. the rate of inflation was always underestimated, total output was usually overestimated and the investigation of the standard of living was not very popular. Inequality measures and measures of wealth were seldom used in practical scientific research.

4. Statistical education at the pre-university level

There is no uniform statistical education at the pre-university level. A big problem is the lack of a so-called statistical culture or statistical literacy. It is related to the use of numbers and their interpretation in everyday life.

As far as the secondary schools are concerned, statistical education coverage varies from rather good statistical theory in mathematical grammar schools to very poor descriptive statistics in some other schools. Because of this variation, students enter university education with very different backgrounds in statistical education. This usually causes many problems in the teaching of statistics at the university level.

5. Statistical education at the university level

Those problems in statistical education from secondary schools are reflected in university statistical education. It is very difficult to overcome these differences in the statistical course which is very rigorous one. Since the knowledge of those entering university is not standard, one must start from the very beginning, dealing with elementary and descriptive statistics in the first academic year, and only later on teaching more advanced statistics.

Statistical courses are now being taught in almost all faculties at the university level, because statistical methodology has been accepted as the most important tool in scientific research. Several years ago, for example, medical statistics was neglected and treated as something not very important. Now, however, there is hardly any research in medicine which does not make heavy use of statistical methods and applications. In fact, though, there is a further problem in universities because the level of teaching statistics varies a great deal between the faculties, from only descriptive statistics for the law faculty to mathematical statistics at a high level for students of mathematics. Between those levels are the others, e.g. the faculty of economics with several statistical courses in statistical methodology and economic statistics, the faculty of medicine with medical statistics at a moderate level, etc.

As the preparation for studying statistics at university level is not so good in secondary schools, students face lot of difficulties if they are to pass the examinations in their statistical courses. For example, at the faculty of economics only 21% of all students pass the exam at their first attempt. This also results from their having insufficient knowledge in the mathematical theory necessary for understanding the statistical theory. Students can try to pass the exam as many times as they want (!) so the bulk of them do eventually pass the examination, albeit after several attempts. However, the Education Bill is going to be changed to enable the study process to be made more efficient. This will mean that students will only be able to make four attempts to pass the same examination, which may have serious implications for students taking statistics.

6. Development of computational statistics

The spread of statistical education among youngsters is connected very much with the expansion of computer science. Students become more familiar with statistics if they evaluate their own examples through computer programs. It is a very good tool for making statistical courses more interesting. It should be emphasised, however, that it can be counterproductive to introduce students to the use of computers before they have been given an in-depth explanation of statistical theory and methods.

The University Computing Center in Zagreb has organised two international conferencies in ITI (International Technology Interface) in Croatia, and is organising the next one, so we can say that lot of things have been done in computational statistics. The shortage of money due to the war situation in Croatia could be the only reason for slowing the development of computational statistics.

STATISTICAL EDUCATION PROGRAMME IN PAKISTAN

Susan Anne Starkings
South Bank University
Borough Road, London, England

1. Syllabuses and examinations

In Pakistan examinations are governed by the relevant University, in Kinnards' case the University of the Punjab. Although this is very similar to what happens in many developed countries the way in which the examination questions are written and assessed is extremely different. The University of the Punjab is concerned that the existing syllabuses were formulated in 1962 and very few attempts to update the syllabi during the last few decades have been attempted. Those modifications that have been proposed (University of the Punjab, 1993) "were abandoned due to some technicalities".

The existing courses in statistics are subject to certain criticism from the University advisors namely:

(i) The courses seem to be lacking in coordination and the students do not fully develop statistical concepts.
(ii) The courses do not appear to be well structured and organised, in relation to the objectives of the syllabus.
(iii) The courses are not only weak in content but are also uneven in distribution of the topics.
(iv) Statistical inference which should form the major portion of the course is not given due importance. It is introduced sufficiently late and most of the teachers cover it hastily.
(v) They provide little insight into the variety and usefulness of statistics.
(vi) They are too technique oriented which over-emphasises computations and under-emphasises the fundamental ideas underlying the statistical reasoning.
(vii) They fail to differentiate between random and non-random phenomena.
(viii) They fail to cope with growing developments in data analysis.

It was felt that there was a genuine need to revise the curriculum and this revision has been undertaken in the light of comments received from various educational institutions, working teachers and educational policy. Habibullah (1992) stated that "The prevalent scheme for examinations in statistics does not provide a real test of the students' ability. The questions set in the theory papers are conducive for rote learning to a considerable extent, and the requirement that the students should attempt five questions out of ten allows too much choice".

Statistics is essentially a practical subject and the syllabus should be altered and designed in such a way as to introduce the concepts and methods required to achieve the aims and objectives of teaching the subject and enable students to develop a better understanding of the subject. For the first time a statistics practical is to be carried out as part of the examination with a 15 per cent weighting. Although this is a major step for Pakistan examinations these practicals are very much, at the present moment in time, prescribed (University of the Punjab, 1993).

"The Practical Examination will be carried out in laboratories with the help of coins, dice, cards, random number tables and other such materials. A minimum of 20 practicals should be carried out".

The content of each section is then stated for example (University of the Punjab, 1993): "Between 50 and 100 hundred observations to be obtained from experiments. Various measures of central tendency and of dispersion are to be calculated".

Statistics teachers in the country are finding it very difficult to carry out these practicals since they involve a considerable amount of time; resources such as dice, coins, balls and other such relevant materials are in short supply; the teachers feel they do not have adequate experience in carrying out and assessing this type of work; and that there is a need to teach statistical methods in real life as opposed to just techniques. Association with other countries who are familiar with this type of work is essential for Pakistan statistical education to develop.

The students also sit 2 written papers and the questions are very much calculation oriented. A typical question is of the form stating some values, not always put in context and calculate some statistical measure, an example from Beg and Mirza (1989) is:

> "The reciprocals of certain values of X are 0.004, 0.0625, 0.05, 0.025, 0.02, 0.125, 0.333, 0.0125. Find the arithmetic mean of X".

This type of question is also used throughout the students time studying statistics. A broad statistics curriculum is not only desirable but essential since students do not learn by having facts hammered into them and rote learning of how to answer questions. Pinder (1987) suggests that: "Most rote learning unless accompanied by understanding, does not remain in the memory. To be retained learning must be understood, and skills practised in a variety of ways".

Howe (1984) states that: "The mental activities of individual students form a powerful source of influence on what is actually learned".

It is essential that the statistics syllabuses, curriculum and examinations are altered in such a way as to provide students with a forum for applying and understanding statistics in a variety of contexts.

2. Statistics Teacher Education Programme (STEP)

This programme was launched in the early part of 1992 and so far 3 training sessions have taken place. Kinnard College statistics teachers were the inauguration team which set up the training programme and provide support to other statistics teachers. Each session has focused on a particular element of statistics. The STEP 3 session concentrated on how to devise, conduct and assess statistics practicals. Before any decision is to be made on how to proceed areas of concern need to be looked at. The points listed below are of more global issues central to statistics in general rather than just to the practicals. Areas of concern raised are:

(i) teachers apparent lack of interest and attitude problems;

(ii) lessons tend to be examination oriented and some teachers only teach between 7 to 10 topics out of the 10 that come up on the examination paper;

(iii) there is a basic weakness in the students to be able to perform this type of work: they are used to routine calculations;

(iv) private tuition work has increased and is seen as a social status and a matter of prestige among students and parents. Hence some teachers will not cover all topics but leave these areas to the private tuition classes, often run by the same teacher. A teacher's salary, in Pakistan, is very low and not enough to support a family. Thompson (1973) reiterates this by saying that: "Local conditions produce quite distinct difficulties and opportunities". Home tuition is essential to maintain a reasonable standard of living;

(v) students prefer to do numerical rather than practical questions since they can achieve full mark;

(vi) students have to answer all questions in English but their first language is Urdu;

(vii) collecting data in real life situations can cause problems particularly for female students in Pakistan. Thompson (1973) also mentions: "Cultural problems are mixed up with social ones".

Teachers at the STEP 3 programme made various suggestions to help solve some of their problems. They suggested that the examination system needs to be changed so that more emphasis be placed on practical elements, refresher courses and workshops for existing teachers should be available, changes need to be made to initial teacher training courses and a reasonable salary be paid to teachers so that private tuition is no longer essential to life and hopefully corruption can be eliminated. Starkings (1992) states: "Statistics in educational terms is relatively new and it appears that in education the needs of the students always seem to supersede the presence of suitably qualified and experienced staff who are ready to teach".

The STEP 3 session concentrated on identifying and trying to find ways of solving the above areas of concern.

3. Annual statistics competition

At Kinnaird College a considerable amount of progress has been made over the last few years to try and bring into statistics a practical element. This has been partly achieved by the Inter Collegiate Statistical Competition. The rules of the competition can be summarised as follows:

1. Students may participate this competition either individually or in the form of teams of either 2 or 3 students of the same college;

2. Each student/team should carry out a statistical project as follows:

a) Decide what is it that you want to find out (for example you may wish to get an idea regarding the proportion of students who take tuition);
b) Collect primary data (real, unpublished data) in order to find a reasonable answer to your question;
c) Analyse the collected data and draw a conclusion; state the limits of your conclusion;
d) Present your project in the form of an attractive poster.

It should be noted that tuition in this context means that a student has private tuition in addition to their normal college lessons. Students of Kinnaird College were not allowed to enter the competition since it was felt that the host college should be neutral.

This is the third year of running this type of competition. The first year had only 2 entries from one college, the second had 135 entries from most of the colleges in Lahore and this year 105 entries were from colleges throughout Pakistan. Although the number of entries was less the coverage of the country was far greater. This was largely due to the tireless effort of the statistics department at Kinnaird College. The entries varied in type of survey carried out. Example titles were "Population of Family re: Education Level", "Why do people take exercise", "Liking or Disliking of Dish Antenna", "Crank Calls", "Smoking and Spirits", "On what basis do people cast vote", to name but a few.

It is surprising how many of the topics appeared to be very western in nature. The competition was on display for teachers and students of any college to come and visit. Several colleges visited with their students and obtained useful ideas to enhance their teaching and to promote further interest in the annual competition. The competition posters are now to be sent to any college in Pakistan that requests them to be displayed. It is hoped that this will encourage many more students to enter the competition in future years. Examples of positive comments taken on the day are 'the competition is both educational and interesting, it brings practical confidence in participants', 'the exhibition is very good attempt and I think it should be continued so that other new coming students can participate in next years competition'. Some problems raised by the teachers are that students have problems defining and applying simple concepts. Present text-books were not really appropriate for teaching practical statistics and that college teachers are not given any training on how to teach. Many statistics teachers reading this may well be in agreement with this in their own country.

Working in an educational system that is predominantly textbook oriented this initiative was to try out new methods for effecting change. The competition certainly has the statistics teachers talking to each other about ways of encouraging their students to be involved in practical work. By providing a forum whereby teachers can discuss ideas and ways of improving their teaching can only be of mutual benefit to the students and the subject progression in Pakistan. It is hoped that this competition will continue to grow and that practical work in statistics will become an integral part of the subject. If this is achieved then the philosophy behind the introduction of the competition will have been fulfilled.

4. Conclusion

There is no panacea for solving the above concerns but help from other countries and international organisations such the ISI and IASE is essential to the growth and future development of statistics in Pakistan. Obviously changes cannot be made overnight and that, although many negative issues have been raised in this paper, there is a considerable amount of good statistics teaching being carried out in Pakistan. Many teachers face considerable financial hardship by entering the profession and are tireless in their pursuit of educating the students to the best of their ability. The pace of innovation has to be constantly monitored otherwise if it is too rapid teachers may be bewildered by new techniques and ideas behind them. The new developments are exciting and refreshing after the more mechanical and repetitive calculations, with the emphasis placed on practical exercises rather than abstract routines, there is a good chance that statistics will come to help students in their ordinary lives.

Bibliography

Beg A. and Mirza M. (1989), *Introduction to Statistics*, The Caravan Book House Publishing Company, Lahore, Pakistan.

Habibullah S. N. (1992), Statistical Education in Developing Countries, The Pakistani Scene, in L. Pereira-Mendoza (ed.), *Proceedings of the International Statistical Institute's Round Table Conference*, Lennoxville, Canada 1992, International Statistical Institute, Voorburg.

Howe M. (1984), *A Teachers' Guide to the Psychology of Learning*, Basil Blackwell Inc., NewYork, USA.

Pinder R. (1987), *Why don't teachers teach like they used to?*, Hillary Shipman Limited, London.

Starkings S. A. (1992), Introducing Data Analysis in the Schools Who Should Teach It?, in L. Pereira-Mendoza (ed.), *Proceedings of the International Statistical Institute's Round Table Conference*, Lennoxville, Canada 1992, International Statistical Institute, Voorburg.

Thompson B. (1973), *Learning to Teach*, Sedwick & Jackson, London.

University of the Punjab (1993), *Syllabi and Courses of Reading Statistics FA/FSc classes XI and XII intermediate examinations 1993*, University Press, Lahore, Pakistan.

ISSUES IN THE TEACHING OF PROBABILITY AND STATISTICS

Contributed papers

Gabriella M. Belli
Manfred G. Borovcnik
Erdinç Çakiroğlu *et al.*
Flavia Jolliffe
Peter Nüesch
Mary Rouncefield
Eric R. Sowey
Vladislav V. Shvyrkov
John Taffe

STUDENT VIEWS ON EFFECTIVE AND INEFFECTIVE LEARNING IN GRADUATE LEVEL APPLIED STATISTICS COURSES

Gabriella M. Belli
Virginia Polytechnic Institute and State University
2990 Telestar Court, Falls Church VA 22042, USA

1. Introduction

Applied or service statistics courses are often viewed by students with something less than enthusiasm. Instructors of such courses have argued for the need to alter attitudes and provide more emphasis on real-world problems and issues using real data and computers (e.g., Blalock, 1987; Roberts, 1987; Singer and Willett, 1990). In a recent article, Snee (1993) reiterated this plea and stipulated that people process information in different ways. Therefore, having different learning styles, they need to be taught with different methods. He concluded by stating that this would result in more favorable attitudes, as well as in greater use of statistical thinking.

Graduate students are adult learners in school by choice, with definite goals for their education, and often motivated by specific job demands. Unfortunately, my 10 years of experience teaching such students is that many of them are enrolled in graduate level statistics courses only because it is a requirement for the degree. For many there is little or no motivation or desire to learn the content or utilize statistical tools. Once students overcome negative attitudes toward statistics, begin to see its value and to think in statistical terms, they do tend to approach the subject differently and some go on to advanced courses, managing to do so even in the face of poor prior background.

What is it that makes the difference for students? What is their perspective on what benefitted or hindered their learning in a first statistics class? Additionally, can students' reasons be categorized based on different student characteristics or learning styles? To answer these questions, the present investigation is a qualitative follow-up to a quantitative analysis that was aimed at determining relationships among learning styles, academic programs, background variables and attitude toward statistics (Belli, 1993). Findings from that study of 155 graduate students indicated that they tended to differ in preferences for how they function cognitively, as well as to differ widely in their readiness for statistics. These differences were found to be somewhat related to mathematics and computer background, attitude and gender.

At the end of the structured questionnaire used for those analyses, students were asked two open-ended questions: (1) What was it that allowed you to learn as effectively as you could? (2) What was it that prevented you from learning as effectively as you could? A content analysis of their written responses was conducted to allow for several comparisons: (1) among students with different learning style preferences, (2) across different disciplines and (3) between males and females.

2. Learning styles

People's preferences or type may be characterized in four bi-polar dimensions based on Jung's theory of psychological type (Myers and Myers, 1980). Two dimensions identify different preferences for using the mind. These preferences relate to: (1) the perceptive process people use to gather information and (2) the decision-making or judgement process people use to reach conclusions. Perceiving may be either a sensing (S) or intuitive (N) process. S relates to practical facts and experience and N deals with possibilities and relationships. Decision-making may be either an objective thinking (T) or a subjective feeling (F) process. These latter terms do not refer to a cognitive-emotional distinction, but merely to a preference for reaching conclusions through impersonal analysis or based on subjective or personal involvement (Lawrence, 1982). By crossing the two dimensions, four basic mental functions that relate to learning style are identified (ST, SF, NT and NF). These were measured using the Kiersey Temperament Sorter (Kiersey and Bates, 1974).

Two additional dimensions in type identification relate to how one uses the perception or judgement process. The extroversion-introversion (EI) dimension indicates a preference for these processes to be outwardly or inwardly directed. The judgement-perception (JP) dimension indicates a preference for dealing with the outer world and dictates whether the SN (perception) or TF (judgement) dimension is the dominant process. J-types, with a proclivity to decision-making, prefer closure and are outcome oriented. P-types, in keeping with an interest toward gathering information, prefer to keep options open and are more process oriented.

3. Student characteristics

The data were collected from students in required entry level statistics courses in education, business and engineering at VPI & SU. The classes were conducted at Virginia Tech's Graduate Center located near Washington, D.C. and were taught by either a resident faculty member (myself) or one of five adjunct statistics faculty during the past three years. Almost all were employed full time and were in either a masters or a doctoral program. Of the original 155 students, 134 responded to at least one of the open-ended questions.

As may be seen in Table 1, most were from education (63%) and over half were female (57%). While over two thirds of the education students were females (69%), males were in the majority in engineering (64%) and business (59%). Students were about evenly split in the SN perception dimension (53% Ss). The majority of education students, however, were Ns. With respect to the TF decision-making dimension, 66% were Ts, as were almost all the engineers (82%). Additionally, students were equally likely to be an E or an I, but almost all were J-types.

Considering the joint distribution of the two personality type dimensions, Table 2 indicates that almost half the students were STs and only 7% were SFs. Males were predominantly STs, while females were split across three categories (excluding SF). Overall, males were predominantly S-types (76%), while females tended to be N-types (64%). Three quarters of the males were Ts as were 58% of the females. This sample differs from the general population where the TF dimension is the only one that is gender related, with 60% of males being Ts and 60% of females being Fs (Lawrence, 1982). The distributions across mental functions by discipline seems to be primarily a reflection of the gender distribution.

4. Aids to effective learning

Responses to the first question were amazingly diverse, from a mere listing of single items to a full page of writing with multiple comments. In the subsequent discussion, percentages for categories of responses are based on the student as the unit of analysis and not the total number of individual comments. Most of the reasons for effective learning fell into four broad categories relating to the instructor, the student, practical applications, and outside assistance. Ten percent of the students provided no response and four said that nothing helped them, as they did not learn effectively. There was much commonality among the comments of students, irrespective of learning style, academic area, or gender. When combined with the skewed nature of the sample and the under representation in some cells, it was difficult to accurately assess group differences. Therefore, the discussion will be more general and only provide comparisons where appropriate.

Table 1. *Gender and personality characteristics of the sample by academic discipline*

Column % within pairs	Education 84 (63%)		Engineering 28 (21%)		Business 22 (16%)		TOTAL 134	
GENDER								
Male	26	29%	18	64%	13	59%	55	41%
Female	58	69%	10	36%	9	41%	77	58%
unknown	2	2%					2	1%
PERSONALITY TYPE								
Sensing (S)	39	46%	18	64%	14	64%	71	53%
iNtuition (N)	45	54%	10	36%	8	36%	63	47%
Thinking (T)	50	60%	23	82%	15	68%	88	66%
Feeling (F)	34	40%	5	18%	7	32%	46	34%

Table 2. *Distribution of gender and discipline according to mental function*

Row percents	ST 61	46%	SF 10	7%	NT 27	20%	NF 36	27%	TOTAL 134
GENDER									
Male	38	69%	4	7%	4	7%	9	17%	55
Female	22	28%	6	8%	23	30%	26	34%	77
unknown	1						1		2
DISCIPLINE									
Education	32	38%	7	8%	18	22%	27	32%	84
Engineering	17	61%	1	4%	6	21%	4	14%	28
Business	12	54%	2	9%	3	14%	5	23%	22

Instructor

Over half of the students either cited the instructor specifically or mentioned instructor generated or selected materials. About half of the instructor comments specified his or her clarity of explanations. For example, that the instructor "wrote information on the board and clarified as we went along" and that "important areas were highlighted". Most of the remaining comments were about the instructor's concern and understanding, apparently very important for some students.

A typical characteristic for S-types, who are linear learners, is a preference for beginning with the known and going systematically, tying new facts to past experiences. They tend to like organized instruction with a detailing of all the steps involved (Lawrence, 1982). One S typified this by saying "the teacher categorized and explained the information in an organized way". However, three other comments about the instructor breaking down material "from simplest to complex" were given by N-type students. Such people, theoretically, are global learners whose intuition provides insight into complexity and who are impatient with details (Lawrence, 1982).

About one quarter of the comments dealt with the instructor's course materials. In particular, students valued handouts that "reinforced the material" and "went thoroughly through each step of the problem". Only nine people mentioned the text.

Students

Of the 41% who said they were a major reason for their own success, three quarters cited effort, doing homework, and extensive study. As typified by one student, the acknowledgement that "spending adequate time on weekly assignments was absolutely essential" was deemed as a very important aspect of effective learning.

Only three students gave the reason for their success as being their aptitude or mathematical bent, another three cited a predisposition or interest in the subject, and seven stated a strong mathematical background. Of interest is the fact that 10 of these 13 statements came from females (mostly in education). This goes counter to expectation, particularly given the prior analysis of 155 students that indicated both a stronger background and more positive attitudes for males, particularly for males in engineering (Belli, 1993).

Practical applications

Although, as mentioned earlier, the literature is rife with the importance of real world examples and practical applications, less than one quarter of the students claimed this was a benefit to their learning. Comments took two forms. The first related to the utility and interest of real-life examples and problems, particularly as they applied to "real, day to day work" and were "immediately applicable in my job assignment". It is typical for S-types to like practical applications and these comments came entirely from STs, most of whom were male engineers. The second aspect related to use of computer software for problem solving so that "we did not have to crunch numbers". This was viewed as providing "freedom to experiment and see various ways to test the same numbers on my own without feeling inadequate". Both aspects were also mentioned together, indicating the utility of the "practical application of statistical software to real world problems".

Outside assistance

The final area, cited by less than one fifth of the students, was outside assistance from a combination of sources. Working with a group was highly valued by these students. One comment confirmed anecdotal reports from students in many classes, i.e., that "we had one member of our group who understood the material and we felt comfortable asking her questions". Although the instructor was the most frequently cited reason for effective learning, there seems to be an apparent pervasive fear of appearing dumb and of not being willing to ask questions in class or of the instructor. As one student put it, in response to the second question, "An uncomfortableness with my confusion prevented me from going to the instructor for help". Hence, the added value of group interaction. Other sources of aid came from families, both in terms of actual tutoring in statistics or computer utilization and of support and understanding. The latter coming, in part, due to the fact that "there are many lost weekends" for families. Although spouses were a source of both types of support, many computer novices benefitted from help provided by their computer-literate children.

5. Impediments to effective learning

Student responses to what prevented them from learning dealt predominantly with time related issues, the instructor, themselves, or course difficulty.

Time related constraints

One third of the students cited various interrelated reasons for lack of time as a primary deterrent to learning. Over half of the reasons consisted of job and family constraints, which were often cited together. Job requirements sometimes caused students to miss class or simply to be too exhausted to concentrate. Family and other responsibilities and commitments also placed a heavy burden on them. As one student said, "generally time to study affected my grade and learning ability".

A quarter of the time related comments focused on the pace of the course and the belief that there was "too much material in too short a time span". A real sense of frustration was evident due to the "necessity of being exposed to different concepts on a weekly basis and not having enough time to fully understand one topic before being exposed to a new one".

Time related complaints were voiced by students in all three disciplines and learning style preference categories.

Instructor

Although the instructor ranked high for half the students, about one quarter felt that he or she was the reason for their inability to learn effectively. This was primarily due to a lack of good communication skills and to not providing good structure and organization for the class. An additional 11% of the students blamed the text and instructor's handouts and materials as being confusing and of little benefit. One student captured what others were saying: "Lack of structure - we should have had: (1) objectives stated, (2) definition of terms, (3) when and why a specified test was used, (4) examples, (5) closure". This comment, which would be typical of an S-type person, came from an NF. The remaining such comments cut across all types. This may reflect the preponderance of J-types, who like to have a plan and know where they are headed. The desire for organization, structure, and handouts that provide step-by-step detail is also likely related to students' time demands as they do not have the time to deal with trying to sort out concepts and information. Hence, they may expect that to a greater degree from the instructor.

It should be noted that fully two-thirds of these complaints were aimed at one adjunct professor in an educational statistics class. In fairness to the instructor, six female students had positive comments in response to the first question. Two, in particular, felt she "worked to reduce anxiety ... was very affirming" and "was not a drill sergeant; did not teach by intimidation".

Students

Eighteen percent of the students blamed themselves for their lack of learning. Over half of these comments dealt with anxiety and fear of the subject, sometimes tied to a poor math background. Comments were both general ("Unfamiliarity with the subject at onset and lack of background") and specific ("Fear and insecurity with the subject matter, which I believe is due to an incredibly bad ninth grade algebra experience"). A few students simply blamed their "own lameness" and "lack of adequate preparation".

Given the fact that these students, on average, have not had a math course in 13 years (10 years with a sd of 7.7 for males, and 15 years with a sd of 10.4 for females), it is interesting to note that only four education students mentioned this specifically. Another education student, however, noted that "the make-up of the class was diverse - from knowing little about math/stat to 'experts'" and claimed that this made it "difficult to ask any but the most essential questions".

Course difficulty and relevance

About 10% of the students found the use of computers to be an added difficulty that was detrimental to their learning. In part, lack of time to learn how to use the computer and software was the problem. But, apparently more important was the lack of adequate instruction and lab assistance and the need to learn both course material and software simultaneously, which "created anxiety and frustration".

A few students felt the course was too difficult or not relevant and that "there was too much emphasis on mathematical techniques and language of a technical nature". Yet, only six students felt a need for more practical applications and interpretation of printouts and articles and one who "would have preferred information to be explained in more psychological/behavioral terms".

6. Discussion

Although the hoped for results of categorizing students' comments about effective and ineffective learning by student characteristics and learning styles did not attain, the information provided by the students dramatizes the problems and adversity faced by working adults in a rigorous graduate program. Responses to both questions showed an incredible amount of consistency among these students. The underlying reason for this may be that the common denominator of being older and already involved in a full-time job (often a long-term career), with many responsibilities beyond schooling, overrides any differences that might otherwise prevail. A similar study conducted with undergraduate, full-time students would assist in determining the viability of this hypothesis and provide an interesting comparison.

An alternative perspective, based on the fact that there are different approaches to measuring learning style, would be to undertake more comprehensive studies using various instruments. Smith (1982) provides 17 such inventories and tests. It is possible that, with a different instrument, students would be classified in ways that would be more consistent with their comments and that patterns would be evident.

Finally, different educational settings may require different learning theories (Dubin and Okin, 1973). A service statistics class populated by at least some non-mathematical adults who have, of necessity, only a part-time investment in their education is a different environment, particularly in an off-campus setting (Belli and Seaver, 1989). Although the field of adult education is primarily aimed at informal education, the focus is the adult learner and some of their work should be applicable. Merriam (1988) proposed a framework for organizing and considering the vast body of adult education literature. Areas such as cognitive style, motivation, and experience seem particularly relevant to understanding and improving the experience in graduate applied statistics classes. Situational characteristics such as part-time v. full-time study and voluntary v. compulsory participation considered in Cross' (1981) Characteristics of Adults as Learners (CAL) theory of adult education could also provide some fruitful directions for further inquiry.

Bibliography

Belli G.M. (1993), *Student type, statistics attitude, and course preference in graduate-level applied statistics courses*, Unpublished paper presented at the annual meetings of the American Educational Research Association, Atlanta, GA, April 12-16.

Belli G.M. and Seaver W.L. (1989), Graduate statistics courses in part-time off-campus programs, *The American Statistician*, 43(2), pp. 86-90.

Blalock H.M. (1987), Some general goals in teaching statistics, *Teaching Sociology*, 15, pp. 164-172.

Cross K.P. (1981), *Adults as Learners*, Jossey Bass, San Francisco.

Dubin S.S. and Okin M. (1973), Implications of learning theories for adult education, *Adult Education*, 24(1), pp. 3-19.

Keirsey D. and Bates M. (1984), *Please Understand Me: Character & Temperament Types*, Prometheus Nemesis Book Company Del Mar, CA.

Lawrence G. (1982), *People Types and Tiger Stripes: A Practical Guide to Learning Styles* (2nd edition), Center for Applications of Psychological Type, Inc. Gainsville, FL.

Merriam S.B. (1988), Finding your way through the maze: A guide to the literature in adult learning, *Lifelong Learning: An Omnibus of Practice and Research*, 11(6), pp. 4-7.

Myers I.B. and Myers P.B. (1980), *Gifts Differing*, Consulting Psychologists Press, Inc., Palo Alto, CA.

Roberts H.V. (1978), Data analysis for managers, *The American Statistician*, 41(4), pp. 270-278.

Singer J.D. and Willett J.B. (1990), Improving the teaching of applied statistics: Putting the data back into data analysis, *The American Statistician*, 44(3), pp. 223-230.

Smith R.M. (1982), *Learning How to Learn: Applied Theory for Adults*, Cambridge Adult Education, Englewood Cliffs, NJ.

Snee R.D. (1993), What's missing in statistical education?, *The American Statistician*, 47(2), pp. 149-154.

INTUITIVE STRATEGIES FOR TEACHING STATISTICS

Manfred G. Borovcnik
Department of Mathematics, University of Klagenfurt
Sterneckstraße 17, A-9020 Klagenfurt, Austria

1. Trends in teaching stochastics

Statistical inference is a methodology which allows one to generalize findings from data. Thereby data can relate to a small subset, a sample of a bigger population; or they can relate to a data generating process which has been observed for a specific period of time. In both cases it is probability models and the concept of random sample which guarantees that the generalization can be done with some accuracy and with some degree of uncertainty. This is why statistical inference has been entirely based on probability theory, also in teaching - at least in France and in German speaking countries. However, probability is mathematically oriented, difficult, and it takes a long time to learn.

Simplifying probability theory. The increasing demand for skills to apply statistics required this problem to be addressed. Thus in the early 1980s there was an endeavour to simplify access to inference. i) By simplifying probability theorems through simulation, e.g. Watkins (1981) or Travers (1981). Probability is not developed as a mathematical theory for the student; instead those relations needed for inference are motivated by simulating the underlying assumptions. This enables a direct approach from descriptive statistics to inference. ii) By reducing probability theory to samples from a finite population, the nonparametric approach from Noether (1973). Probability is not developed as a theory for the student. The given data are multiplied by systematic permutations to yield the population. The inference then is based on whether the value of the test statistic for the given data is among the extreme values of this statistic for all the possible permutations. iii) By motivating methods through the context to which they are to apply, e.g. Holmes (1983) or Kapadia (1983). Many methods comply with common sense when they are bound to a specific subject matter rather than taught as abstract procedures.

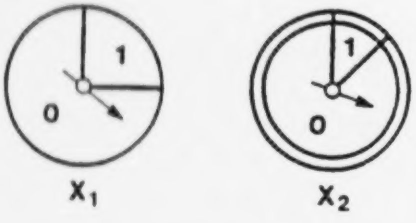
Reduction to data analysis. The late 1980s are signified by a movement back to basics, back to data analysis. The aim of this strand was to eliminate probability theory completely, at least in the earlier stages of teaching. It developed a new type of generalizing results from data which is not based on the random sample argument. The development goes back to Tukey's Exploratory Data Analysis (1977) which investigates data without a theory of probability. Herein, data are analyzed in a detective manner, from various angles, by multiple methods and in an interactive way between the results of intermediate analyses and the analyst's knowledge of the subject matter from which the data originate. Bibby (1983) and the 'Statistics in Society' course from the Open University were first to introduce this approach into Teaching Statistics. The Quantitative Literacy Project combined the exploratory approach with simulation ideas from 1986 onwards (Scheaffer, 1991). Recent developments, however, reduce this approach more and more to exploring data for its own sake: Gordon and Gordon (1992) summarize the state of the art.

Stochastic thinking. There is nothing to say against the exploratory approach; however, there is much to argue against reducing probability and statistics to it and neglecting the probability part of the whole. The psychologists Falk and Konold are the only critical voice in Gordon and Gordon (1992) who remind us that probability establishes a distinct approach towards thinking about reality. This approach is different from logical and causal thinking; it is difficult and at the same time important. Therefore teaching should clarify how it could contribute to solve real problems. The author will focus upon showing why this view on reality is important and how one can teach it successfully.

2. Specific problems of stochastics and its teaching

No learning by trial and error. Trial and error is a fundamental type of learning which accompanies us from our earliest days. This kind of learning is also true for mathematical concepts, even for the more difficult ones. Some theories of learning have integrated this fact. Piaget's operational stage of learning precedes the formal stage of understanding; more refined theories develop a mutual interplay between operations and reflections, see e.g. Dörfler (1984). According to this approach operations upon a material representation of the concept allow for feedback to the individual to revise his intuitions if necessary. Thus more refined forms of the concept may be established in the individual's mind. For example, the notion of number can be established from such concrete operations. It is not until very sophisticated levels that counterintuitive features emerge which are not based on concrete operations. Among such relations are: i) subsets can have the same cardinality as the reference set; ii) the rational numbers are countably infinite even if they have no minimal distance between two numbers and thus are not suitable for counting.

It is the author's thesis that this kind of learning is not possible for probability concepts even from the very basics. Thus, this phenomenon establishes the difficulties which subjects encounter with learning probability. An example similar to one from Falk and Konold (1992) will illustrate the fundamental problem that probability concepts reveal features which allow no direct feedback. In the diagram below are two spinners. The player has to choose one of the spinners and wins if the pointer will stop in the sector marked with a 1. Which wheel of fortune is the better choice? This simple question linked to the very basis of probability has no simple answer in terms of learning by trial and error. The choice of the worse spinner 2 may lead to success whereas the choice of the better spinner 1 may lead to failure. The latter event, however, immediately causes the player to speculate about the 'reasons' behind the failure and may obstruct progress in concept acquisition. Children are led astray by the special margin of spinner 2 (at the age of 7, this author's own experiment). The popularity of astrological predictions shows that such behaviour is not bound to children.



Clearly, longer series of spinning would allow for the feedback needed for concept acquisition. However, such series are usually not the focus of our experience. Furthermore, problems turn up again when one tries to generalize insight from repeated experiments to a single new outcome. It remains an open question how to supply effective experiences from which individuals can learn from trial and error that probability is a concept useful to apply to a single case because, basically, probability is a theoretical and not an experimental concept (see also Steinbring, 1991).

The gap between actions and reflections. There are various approaches to mathematics education which are based on an interplay between actions and reflections starting with Piaget; a modern version being Dörfler

(1984). These may be interpreted as reconstructing the learning process as intelligent trial and error. Actions upon a material representation of the concept in question form the ingredient of a process of reflection during which new objects will emerge from the concrete operations. Hierarchical steps of action - reflection - new object lead to the concept acquisition in an increasingly complex form.

The following simple example shows that probability notions are lacking in such an operational understanding. With coin tossing, the operation is predicting the outcome of the next toss whereas reflection means to evaluate the weight of heads (or tails). The result of the reflection process may yield the relative weights of 1:1, or the probability of 1/2 for heads. However, this result is of no use, at least no direct use for the prediction problem as the operation necessitates an exact and certain prognosis of the next outcome. For empirical evidence of this intuitive conflict see Falk and Konold (1992). Actions and reflections are based on different levels and the feedback between these different levels is missing for probability.

The basic difficulties in the coin tossing context may be best described by a mutual interplay between individual intuitions and theoretical demands. The individual witnesses an intuitive conflict right from the beginning as on the one hand he feels incapable of predicting the next outcome with absolute certainty and on the other hand he deeply desires control of that chaotic situation. At this point mathematics enters the stage and promises to calculate randomness. Any result of this theory will in due course be re-interpreted as a guide for action. Consequently, mathematical statements are either overinterpreted (like the law of large numbers used to derive the prediction of tails after five consecutive heads) or abandoned, reestablishing e.g. causal schemes for the prediction; see Falk and Konold (1992) for their dominance.

Mathematics offers a justification but yields no insight. Conceptual thinking can neither be reduced to mathematics nor to its applications. To *justify* concepts necessitates modes of thought different from those required to *understand* the concept. This might not be a problem as long as one remains within pure mathematics but things are different when one applies or teaches mathematics. Then it has to be brought into the open that a logically consistent argument to justify a concept cannot mirror all of its features. The concept of independence, for example, is mathematically reduced to the multiplication formula and becomes the basic ingredient of theorems about repeated random experiments like the law of large numbers or the central limit theorem. Independence thus gains an important role within theory, but its mathematical definition by no means affects the causal ideas individuals relate to it. In this context, Freudenthal (1983) used the phrase of-mathematics as turning the things upside down.

Furthermore, the axiomatic structure of probability is not unique. The classical approach is based on the tendency of an experiment to produce stable relative frequencies whereas the Bayesian approach centres the theory around the idea of probabilistic judgements in the form of a weight of evidence. Thus, different axiomatic theories justify different intuitive ideas and do not cover the ideas of the other position quite well. Conditional probability for example, has a poor representation within classical probability. Now many of the historical puzzles and paradoxes are signified by an intermixture of intuitive ideas of both approaches. While it may be an interesting scientific programme to develop a selfcontained closed theory for one idea of probability, this programme does not yield insight into the true character of probability which remains a mixture of different ideas. With this in mind, Barnett (1973) tries to reconcile the competing positions of classical and Bayesian statistics.

3. Intuitive strategies for teaching statistics

It was argued that an operational understanding is lacking in probability as there is no direct feedback from the level of actions to that of reflections and vice versa. A far more promising approach to teaching statistics is by the interplay between intuitions and theory by Fischbein (1987). According to that, raw primary intuitions are revised by partial theoretical inputs into secondary intuitions whereby intuitions generally form the key to understanding and acceptance of the theory. The approach serves to describe the specific difficulties as well as serving to design teaching activities. For a description of the interplay in the coin tossing context see Borovcnik (1991).

Intuitive reformulation of concepts. Inner mathematical features of a concept are only one side of concept learning. Relating these features to intuitive ideas of individuals will help them to *understand* and *accept* theoretical results which is crucial for teaching to be effective. Some key concepts now double the difficulties as there are inadequate related intuitions *and* their mathematical representation is streamlined to a specific form according to a theory which is far from being intuitive. These difficulties will be illustrated by the Falk problem for the concept of independence and dependence. For intuitive difficulties with these concepts and the use of this problem in empirical research see Borovcnik and Bentz (1991). There are a lot of related puzzles in which teaching the usual mathematical representation will not effectively overcome the misconceptions, as the public discussion around Marilyn vos Savant and the Player's Dilemma shows (Morgan *et al.*, 1991).

The Falk problem deals with a simple urn with two white and two black balls. The urn is thoroughly shaken, two balls are blindly drawn out of it, one after the other without replacement. Consider two questions:

a) What is the probability that the second ball is white, given that the first ball is white?

b) What is the probability for the first ball being white, given that the second is white and the colour of the first is not known?

In what follows, the event W refers to white balls, B to black balls; the index describes the number of the draw. The usual solution, even if supported by tree diagrams will not change wrong causal intuitions because it is too formal (Borovcnik and Bentz, 1991):

$$P(W_1 \mid W_2) = \frac{P(W_1 \wedge W_2)}{P(W_2)} = \frac{P(W_1) \cdot P(W_2 \mid W_1)}{P(W_2)} = \frac{1}{3}.$$

Intuitively there is a big difference between the timely forward direction and the backward direction. A new urn 'causes' the new probability 1/3 in the forward question a) whereas the backward question b) is lacking in such an obvious urn. How to reformulate the representation to make the answer and the concept intuitively more accessible? Conditional probabilities are simplest to introduce via odds, i.e. relative probabilities or weights. The probability 1/3 e.g. amounts to odds of $1/3:2/3 \equiv 1:2$. A *conditional bet* on the first ball being white after the second is known to be white involves comparing the relative weights for W_1 in the face of the information W_2, i.e.

$$P(W_1 \wedge W_2) : P(B_1 \wedge W_2) = \frac{1}{2} \cdot \frac{1}{3} : \frac{1}{2} \cdot \frac{2}{3} = 1:2.$$

As odds of $a{:}b$ yield the probability $\frac{a}{a+b}$, the conditional probability in question b) is

$$P(W_1 \mid W_2) = \frac{P(W_1 \wedge W_2)}{P(W_1 \wedge W_2) + P(B_1 \wedge W_2)} =$$

$$= \frac{P(W_1 \wedge W_2)}{P(W_2)} = \frac{P(W_1) \cdot P(W_2 \mid W_1)}{P(W_2)}.$$

Now the symmetry between forward and backward question a) and b) is obvious. However, the full potential of this representation of conditional probability will not unfold until the Bayes' formula comes to the fore. In its simplest form Bayes' formula deals with two prior hypotheses H_1 and H_2 with prior probabilities $P(H_i)$, an experimental evidence E with known probabilities $P(E \mid H_i)$ under the two hypotheses. Bayes' formula then gives the posterior probabilities

$$P(H_i \mid E) = \frac{P(H_i) \cdot P(E \mid H_i)}{P(H_1) \cdot P(E \mid H_1) + P(H_2) \cdot P(E \mid H_2)}.$$

With odds this reads as

$$\frac{P(H_1 \mid E)}{P(H_2 \mid E)} = \frac{P(H_1)}{P(H_2)} \cdot \frac{P(E \mid H_1)}{P(E \mid H_2)}$$

posterior odds = prior odds × likelihood ratio.

Now one immediately sees that the posterior odds, the new weighting of the hypotheses are linearly dependent on the prior weighting *and* the relative likelihoods of the hypotheses under the empirical evidence. The odds representation of Bayes' formula shows the direct impact of the various inputs to the final probability which dissolves many a related puzzle. Furthermore, it shows that conditional probabilities are neither bound to time direction, nor to causal influence; they merely reflect an indication in summarizing the new state of knowledge in the form of a new probability. The reason for the usual approach to conditional probabilities being so clumsy is that the approach is done *completely within* classical probability whereas the concept is very natural within the Bayesian probability theory; odds re-establish this natural character of the concept. For further details and teaching consequences see Borovcnik (1992).

Teaching as clinical interview. A clinical interview is designed to reveal an individual's thoughts and thereby his/her comprehension of a mathematical concept. This author believes that such an approach is doomed to failure in its pure form as any question, any diagram, or any social interaction will establish a theoretical input for the individual and thus he will *react* to that. The individual will neither reveal everything of his understanding nor will he remain unaffected by an input. This dynamics, however, is not at all crucial in teaching where an intervention is intentional. The difficulty in teaching is, as was stated already, to connect the theoretical input to the pre-existing intuitive world of the individual. It was also argued, that a direct feedback between the operational and reflective level of a concept is missing, which obstructs learning by trial and error. Any teaching in probability is therefore faced with the problem of establishing links between theoretical demands and individual thought.

If such links are not supplied, the teacher might ask on the reflective level whereas the student answers at the operational level. For example the student might correctly evaluate the probability of 1/2 for heads but as this gives no clue for action, he/she might deliberately choose heads (or tails). If the teacher asks for a justification for that choice, the student might be completely confused because in his/her mind there is no such justification. The confusion finally might lead to a complete breakdown of the communication in class. Interviews might overcome this action reflection conflict.

Another link between theory and intuitions are urns and other games. They are a useful reference point which facilitates communication in class. Students may easily read off numerical probabilities, calibrate their individual scale of probability, and get insight into theoretical relations. Urns may therefore be perceived as partial images of a restricted theory which help to clarify intuitions. It is useful to map a real situation onto a suitable urn to deal with the questions. However, empirical investigations do show that intuitions about urns are by no means uncontroversial even in the simplest cases (Borovcnik and Bentz, 1991). If used as a medium between theory and intuitions, these inadequate associations have to be eliminated first. During this phase of teaching, feedback to students' thoughts is crucial and worthwhile. Teaching in the form of an empirical interview might best help with the multifaceted idiosyncratic ideas.

The following variant of the Falk problem should illustrate the potential of the interview approach to teaching. More than 50% of adult students confronted with the backward question b) give the answer 1/2 and justify it with causal arguments. To confront the inadequate causal scheme in this context *and* to establish the idea of new weighting of possibilities when further information is available, the author confronted students first with a third draw and the event second and third ball being white. As drawing is without replacement, it is then logically impossible for the first ball to be white. A first shock for the time dependent causal thinkers - there might be logical reasons to integrate *later* information into the weighting of *earlier* events. The harder causal thinkers needed more: the urn was changed to 4 white and 4 black balls. Again, first ball is unknown, second ball white - the answer was 1/2, the argument causal; second and third ball white - 1/2 again. The urn was changed to 50 white and 50 black balls and subsequently the following events were presented to the students: second ball white, second and third white, ... second to 49th white, second to 50th ball white. In the last case it is still logically possible for the first ball to be white but even the hardest causal thinkers had then become shaky. They recognized that the later information must have an impact on evaluating the first ball even if it is not logically cogent. The next step was to ask *when* this information began to exist - after 45 subsequent white balls, or after 6, or even when only one white ball in the second draw? Now students saw that the conditional probability $P(W_1 \mid W_2)$ has nothing to do with causal deliberations, but merely reflects the new state of information and that there was new information to express in this problem. This insight was facilitated by the interview teaching as neither students nor teacher would otherwise have known of each other's ideas.

Teaching by analogies. A conceptually difficult situation from probability is mapped onto a situation from another context in which the learner is experienced and can deal with problems and restrictions of the results. This can be extended to a teaching strategy; the potential will be illustrated in the context of betting and error theory.

Historically, the development of central concepts in probability come from studies of games of fortune. It is still possible to sharpen individual ideas of probability using this context. Betting is a side strand of today's curriculum and the fair prize of a game is a derived concept, namely the expected value. A basic difficulty with expected value is that it only makes sense for a long series of games. With such a series, the empirical mean will approximate the expectation. However, it is far from uncontroversial that one can deduce what is the fair prize for a *single* bet from that expectation. With odds, it is much easier. Odds are simply another representation of probabilities. If the odds are 1:2 for the winning condition then the stakes should amount to 2:1 which means that one has the possibility to win $2 for a stake of $1, the stake being lost if the winning condition fails to occur. For further details and more intuitive justifications see Borovcnik (1992).

It is important to state that with odds it is very natural to fix the stakes for a *single* game without solving the problem of predicting the outcome of this game. This might overcome the misconception of many individuals who accept probability only in cases where it predicts the exact outcome with certainty. Odds also serve for a new introduction into conditional probability and the Bayes' formula as was already shown here. The idea of the prize of the game makes various relations quite simple: if one plays two games then the prizes should add which yields the additivity of the expected value. For the binomial model situation of playing n times the same game with success probability of p, the single prize of p adds up n times which gives np for the expected value of the binomial distribution. The advantage of such analogies is that mathematical results may be perceived by the learner in advance and mathematics can focus on the issue of how the intuitive situation has to be made precise and which assumptions still have to be considered.

Error theory within geodesy was the second big stream of development for probability theory; the normal distribution and the method of regression are intimately related to it. The Bernoulli law of large numbers deals with the contraction of the distribution of the (theoretical) relative frequencies towards the unknown probability of the binomial experiment. This is proven by Cebychev's inequality with the help of the variance of the binomial distribution. This variance is tricky to calculate and the result is usually not well motivated, a very bad basis for the interpretation and acceptance of such an important theorem by the learners. There remains in the learner's mind a vague idea that the relative frequencies converge to the underlying probability and this forms the basis of misconceptions. An analogue law of large numbers holds for the mean of repeated trials.

In the context of error theory, this result can be intuitively anticipated. The analogy starts with the repeated measurement of a physical quantity a. Too large and too small numbers will eliminate each other if one calculates the mean of the single measurements which leads to the mean being a much better representation than the single measurement. To refer this to the measurement of an unknown probability p, the single measurements are 1 if the event occurs and 0 if the event fails to occur. Despite unusual values however, the mean coincides with the usual fraction of trials with the event 1, which is the relative frequency. Borovcnik (1992) gives details of how to develop a class experiment to get measurements and to study the precision of the repeated measurement. The longer the series the less scattered are the means (the relative frequencies); the procedure for measuring p tends to get more precise if it is based on more data.

This analogy avoids some of the basic misconceptions about the convergence of relative frequencies as it does not investigate a single development of the frequencies with increasing series; it focuses on the distribution of the means (the frequencies) with a fixed number of measurements and only then compares this distribution for different numbers of data. The growing precision is to be expected within error theory. Probability again gets a meaningful interpretation, without the precise prediction of a single outcome as one investigates the variation of measurement values (means) for different sizes of data. The process of making things more precise will lead to the binomial distribution, the central limit theorem, and confidence intervals. Again, it is important that the analogy has enough potential to prepare a way for learners to anticipate mathematical results. For more details about this and other analogies see Borovcnik (1992).

4. Summary

Probability is more a heuristic than a modelling description of reality. There is no direct feedback from reality which is often so helpful for the acquisition of concepts in other branches of mathematics. This is why intuitive ideas gain such an important role for the teaching and learning of probability. Intuitions can prove an obstacle to the acceptance and the comprehension of the concepts. There is an urgent need for a representation of mathematics which relates it to the ideas of the learners so that they can revise their intuitions wherever this is necessary. This paper illustrated three strategies to help an intuitive understanding of probability.

Bibliography

Barnett V. (1973), *Comparative Statistical Inference,* Wiley, New York.

Bibby J. (1983), EDA for 16-19's, in D.R. Grey, P. Holmes, V. Barnett and G.M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics,* Teaching Statistics Trust, Sheffield, Vol. 1, pp. 277-281.

Borovcnik M. (1991), A Complimentarity between Intuitions and Mathematics, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics,* International Statistical Institute, Voorburg, Vol. 1, pp. 363-369.

Borovcnik M. (1992), *Stochastik im Wechselspiel von Intuitionen und Mathematik,* Bibliographisches Institut, Mannheim.

Borovcnik M. and Bentz H.-J. (1991), Empirical Research in Understanding Probability, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters: Probability in Education,* Kluwer, Dordrecht, pp. 73-106.

Dörfler W. (1984), Actions as a Means for Acquiring Mathematical Concepts, in *Proceedings of the Eighth International Conference on Psychology of Mathematics Education,* Sydney, pp. 172-180.

Falk R. and Konold C. (1992), The Psychology of Learning Probability, in F. Gordon and S. Gordon (eds.), *Statistics for the Twenty-First*

Century, MAA Notes, 26, The Mathematical Association of America, pp. 151-154.

Fischbein E. (1987), *Intuition in Science and Mathematics. An Educational Approach*, Reidel, Dordrecht.

Freudenthal H. (1983), *Didactical Phenomenology of Mathematical Structures*, Reidel, Dordrecht.

Gordon F. and Gordon S. (eds.) (1992), *Statistics for the Twenty-First Century*, MAA Notes 26, The Mathematical Association of America.

Holmes P. (1983), Developing Materials for the Schools Council Project in Statistical Education, in D.R. Grey, P. Holmes, V. Barnett and G.M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics*, Teaching Statistics Trust, Sheffield, Vol. 1, pp. 269-276.

Kapadia R. (1983), A Practical Approach to Statistics, in D.R. Grey, P. Holmes, V. Barnett and G.M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics*, Teaching Statistics Trust, Sheffield, Vol. 1, pp. 169-178.

Kapadia R. and Borovcnik M. (eds.) (1991), *Chance Encounters: Probability in Education*, Kluwer, Dordrecht.

Morgan J.P., Chaganty N.R., Dahiya R.C. and Doviak M.J. (1991), Let's Make a Deal: The Player's Dilemma, *The American Statistician*, 45, N. 4, pp. 284-287.

Noether G. E. (1973), *Introduction to Statistics: A Nonparametric Approach*, Houghton Mifflin Comp., Boston.

Scheaffer R. (1991), The ASA-NCTM Quantitative Literacy Project: An Overview, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 1, pp. 45-49.

Steinbring H. (1991), The Theoretical Nature of Probability in the Classroom, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters: Probability in Education*, Kluwer, Dordrecht, pp. 135-168.

Travers K.J. (1981), Using Monte Carlo Methods to Teach Probability and Statistics, in A.P. Shulte and J.R. Smart, *Teaching Probability and Statistics*, NCTM, Reston VA., pp. 210-219.

Tukey J. W. (1977), *Exploratory Data Analysis*, Addison Wesley, New York.

Watkins A.E. (1981), Monte Carlo Simulation: Probability the Easy Way, in A.P. Shulte and J.R. Smart, *Teaching Probability and Statistics*, NCTM, Reston VA., pp. 203-209.

SECRET CODES IN STATISTICAL EDUCATION

Erdinç Çakiroğlu and Tibor Nemetz
Middle East Technical University
Department of Sci. Education, 06531 Ankara, Turkey

1. Introduction

The utility of solving simple substitutional cryptograms in statistical education at school level has been widely demonstrated (see Nemetz, 1993). In England, for example, national, project type competition was organised in 1989 in this area and it has been proved very successful. For the goals and history of these competitions, the reader is referred to Hawkins (1989). There are other similar simple examples within cryptography which can be fruitfully used for demonstrating the power of statistical methods in "real world" problems. This paper discusses one of them, namely the transpositional ciphers. Their solution is a feasible task from middle school years up to university level. At pre-university level, instead of using likelihood ratios, we propose a different statistics, which matches the age of the children and "near optimal". Connection to computer programming is straightforward, applicable both to easy and to more involved tasks. By working through a few examples, the need for applying statistical methods and the nature of the statistical decisions will become evident.

At university level, this problem offers an excellent opportunity to introduce "real life" hypotheses testing problems and, at the same time, highly theoretical divergence measures probability distributions. For teacher training, the lesson to be learnt is that theoretically based procedures can be transformed to the level of understanding of school children. The relations to graph theory can also be pointed out. Such an abstract notion like the existence of the Hamiltonian circuit in a graph obtains clear meaning.

The primary goal of this presentation is to contribute to the collection of easily solvable project topics for teacher training, which could be accepted by the teachers to be included in the school practice.

2. Transpositional secret codes: an explanation

This encoding is applied to written texts. The message which is to be sent over an unsafe channel is subjected to a coding procedure with the goal of hiding the secret information contained in it. To this end, the transpositional codes divide the message into blocks of fixed length, which we denote for simpler reference by the letter B. Then they transpose the positions within the given blocks by applying the same permutation for the positions in all blocks. The last block of the message is usually not complete. There are different rules to fill it in, e.g. randomly, which are not essential for the present purposes. The transformed blocks are sent over the channel, and the legal receiver simply applies the inverse permutation (known to him) to decode the original message. Usually, however, unwanted eavesdropping is possible, and a third, undesirable party can get hold of the cryptic text. In this paper we escort him on the way of breaking the secret code. For interested readers we recommend Kahn's "Bible of cryptographers" (1967), where the history of such crypto-codes is easy to follow. For your reference, we provide you with an example in the Appendix.

3. Breaking the code: a hypotheses testing problem

Before turning to educational aspects, let us formulate the codebreaker's task in the language of statistics. Say, he obtained N complete blocks, this is his sample. He has exactly $B!$ simple hypotheses, each of them is a possible candidate for the applied permutation. He knows for sure, that one of them was really applied. He also knows the habit of the correspondents, that the actual permutation was randomly chosen, i.e. with equal probabilities, from the set of all possible cases. Therefore, we have the Bayesian Hypothesis-Testing Problem with a large number of simple hypotheses.

At the beginnings, he does not know the underlying probability distributions, but we may suppose that he knows the source, at least in the sense, that he can, in principle, inspect a sample of arbitrary length to obtain a statistical estimation of the underlying probabilities with any desired precision at any given level. But now he runs into the conflict of theory and practice: memory and time constrains do not allow him to carry out this task.

Again, theory gives the key to proceed further. It is well known that the Bayesian error probability for a multiple number of simple hypotheses can be bounded by a linear function of the maximum error probability for the pairs of the constituent hypotheses, moreover the error-exponents in these cases are the same. Therefore he can decompose the original testing problem into a number of not-so-complex problems, without loosing too much, in the sense of optimal error probability. A possible decomposition can be obtained in the following way. To describe it, we introduce some formal notations:

- The original message is composed of blocks $M(i)$, $i = 1,2,\dots,N$, and the *jth* character of the *ith* block is denoted by $M(i,j)$, $j = 1,2,\dots,B$.

- The encoder uses the inverse permutation of P (we use this notation just to simplify formal expressions).

- The coded blocks are denoted by $C(i)$, $i = 1,2,\dots,N$, and *jth* character of the *ith* block is $C(i,j)$.

  Now we formulate B decision problems:

Decision problem $D(K)$, $K=1,2,\dots,B$:

Which is the true Hypothesis:

The *Kth* letters of the coded blocks are followed by their *Jth* letters in the corresponding message blocks, $J = 1,2,\dots,B$, $J<>K$ (i.e. the letters $B(I,P(J))$ are subsequent to $B(I,P(K))$ in $M(I)$),

or

The *Kth* letters of the coded blocks are the last letters in the message blocks.

It is easy to see, that a correct decision on P is equivalent to correct decisions for all $D(K)$. Also, for large sample size (i.e. for large N) the last Hypothesis in $D(K)$ is equivalent with large probability to the claim, that there is no position J, such that all the letters $B(I,P(J))$ are subsequent to $B(I,P(K))$ in $M(I)$.

Furthermore, since P has been chosen uniformly from all possible permutations, the hypotheses in all problems $D(K)$ have the same probability, none of them plays any special role. The maximum likelihood principle tells then that we have to accept the hypothesis J, for which the probability

(1) Pr { $M(I,P(K))$ is followed by $M(I,P(J))$, $I = 1,2,\dots,B$, $J<>K$)} is the maximum, unless the probability

(2) Pr { $M(I,P(K))$ is the last letter in all $M(I)$ } is larger, in which case this is the decision.

We have to turn to experiments to find out what the underlying probability distributions are, and to establish that they are different from the one-step transition distributions are, and this is not a difficult job. Under this condition the Bayesian error probabilities go to zero exponentially as N increases, and the error exponent is the same for all decision problems $D(K)$ (and is the same as for the original Hypothesis testing problem on P); see e.g. Nemetz (1972).

Now we can propose the following decision procedure.

Decision procedure

Step 1. Prepare a conditional frequency distribution $F(Y|X)$ from consecutive characters of a "long" source sequence, which gives an "acceptable" estimation of the one step probability transition distribution, and use this approximation instead of the unknown real one.

Step 2. For any given K, $K = 1,2,\ldots,B$ compute for all $J<>K$, $J = 1,2,\ldots,B$ the negative of the log-likelihood functions

$$(3)\quad -[\log F(C(1,J)\,/\,C(1,K)) + \log F(C(2,J)\,/\,C(2,K)) + \ldots + \log F(C(N,J)\,/\,C(N,K))]/N,$$

and take the value J that minimizes it. Note, that with large probability this is uniquely defined if N is large enough. If not, solve the breaks arbitrarily. Technically, one works with the negative of the likelihood function to get positive values, therefore we use the minimum.

Step 3. Define a directed graph with vertex-set as positions in the blocks, and connect position K to position J with a directed edge if J is the value which minimizes the likelihood functions for the given K.

Step 4. Find a maximum length Hamilton line in the graph above, and use it to define a permutation P which reconstructs the original message box. If such a line is not exhaustive, proceed arbitrarily.

We formulate a number of claims concerning this decision procedure in somewhat loose terms. Most of them is known or easy to prove.

Claim 1. Due to the law of large numbers, the likelihood functions in (3) converge against their expected values. If J yields the correct hypothesis, then these are the differences of the conditional Kullback-Leibler Informational Divergencies $D(F(2,1)||P(2,1))$ and the first order entropy $H(1)$ of the source. Here $P(2,1)$ denotes the "true" probability distribution of the consecutive characters of the (stationary) source sequences. For $D(.||.)$ and its educational oriented introduction we refer to Nemetz (1972). Under mild conditions this difference goes to zero if the size of the sample used to derive the approximation $F(\ldots)$ increases.

Remark: The expected value is infinite, if $F(Y|X) = 0$ whenever $P(Y|X)$ is not. The usual trick to overcome this difficulty is to modify the frequency distributions a little bit by adding 1 to all possible frequency entries before forming the relative frequencies. In what follows we always use $F(.,.)$ with this meaning.

Claim 2: The conditional distributions of the *Kth* successors of a source symbol are monotonically "nearer" to the (stationary) one dimensional distributions. This implies, that if J is not the proper hypothesis, the expected value is larger than in the case of the true J. Therefore the Bayesian error probability of the proposed decision rule goes to zero when N goes to infinity.

Remark: If K is large, practically there is no observable dependence. For our present purposes $K > 5$ is large enough. Our empirical studies show that in case of Hungarian and Turkish written text sources, for blocklength $B = 20$ and message length of $N = 50$ blocks, the error probability is under 1 per cent.

Remark: In case of written language sources, an acceptable approximation $F(.,.)$ can be obtained by analysing a text of around 50.000 characters. Such a job was a slaves' work two decades ago, but now is within the range of middle schools. Historically, language statistics have a long standing, they were used in the last century (see Kaeding, 1898).

4. Tasks for teacher training college students

The appropriate place in the curriculum for solving transpositional codes is the time of discussing hypotheses testing problems. The students may observe the necessity of collecting data, they can feel that the problem is not artificial, and sense the power and the nature of statistical decisions. Also, they can practise the way of applying theoretical results. Performing the tasks in this section may be instructive to social science, business and electrical engineering students, as well.

Task 4.1. Construct a file of a source sequence of about 50.000 characters, and another smaller one playing the usual role of the control group. The first one is to be used to construct a frequency table, the second one is for checking the goodness of the decision rule.

In constructing the source files, work-sharing should be applied. In a class of 25-30 students this means of inputing 2.500 characters per students, which is a realistic demand. It is to be noted, that the output data files can be used for many statistical illustrations, as well.

Task 4.2. Prepare frequency tables of the *k*-step transitions of the source data file, and extend them to contingency tables. (Here *k*-step transition relates to bigrams of the source sequence where the first and the second characters are separated by exactly $k - 1$ symbols inbetween. Care should be taken to explain this). Note, that a single program will do if *k* is treated as a parameter.

Task 4.3. Prepare the log-likelihood tables. For technical reasons, outlined in the previous section, add +1 to all entries in the frequency tables, and divide the entries by their row-sums to get the "forward" *k*-step transition relative frequencies, by column-sums to get backwards transition relative frequencies, and by the total-sum to get the absolute relative frequency tables. Note that all relative frequencies appearing here are strictly positive. Thus it is meaningful to take logarithms of the entries. The resulting tables all called log-likelihood tables.

```
                                                   ____________________________
0.0009   1.70 - 1.89   |                         |                            |
0.0000   1.80 - 1.99   |                         | MIN=1.7435;   MAX=3.3052   |
0.0062   1.90 - 2.09   | X                       | MEAN=2.5343                |
0.0214   2.00 - 2.19   | XXXXX                   | VAR =0.0530                |
0.0375   2.10 - 2.29   | XXXXXXXXX               |____________________________|
0.0830   2.20 - 2.39 . | XXXXXXXXXXXXXXXXXXXX
0.1320   2.30 - 2.49   | XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
0.1588   2.40 - 2.59   | XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
0.1945   2;50 - 2.69   | XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
0.1427   2.60 - 2.79   | XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
0.1008   2.70 - 2.89   | XXXXXXXXXXXXXXXXXXXXXXXXX
0.0607   2.80 - 2.99   | XXXXXXXXXXXXXXX      ____________________________
0.0277   2.90 - 3.09   | XXXXXX              | Distribution of the values |
0.0241   3.00 - 3.19   | XXXXXX              |    of statistics (3)       |
0.0045   3.10 - 3.29   | X                   |   for written Turkish      |
0.0045   3.20 - 3.39   | X                   |  with sample size of 1150  |
0.0009   3.30 - 3.49   |                     |____________________________|
```

Figure 1. *Distribution of statistics (3) under the null hypothesis*

Task 4.4. Prove that the goodness-characteristics of the decision procedure do not depend on the choice of the permutation *P*. Conclude that therefore any performance analysis can be carried out using the identity permutation, only.

Task 4.5. Take a sample ($(U(i), V(i))$) of pairs of characters from the "control" text, which are separated by exactly $k-1$ letters, for $k=1, \ldots, 20$. Determine the empirical distribution function of the statistic $-\log F(V,U)$. Find a graphical representation that can convincingly show, why the statistics (3) works. Fig. 1 shows what kind of graphs are to be expected.

Task 4.6. Take $B=20$, $N=50$, $P=$ identity, and use the control text to analyse the performance of the decision rule empirically.

At that point we stop listing further tasks, but, obviously, every experienced instructor can supplement them in a number of ways. Now we formulate two research problems.

Open problem 1. Construct an appropriate source distribution model (e.g. Markov chain) and determine the probability, that the decision rules above define a complete Hamilton line.

Open problem 2. Generalize the decision rule to include the first-step backwards transition probabilities, which could define "backwards edges" in the position graph, and investigate if it improves the performance characteristics.

5. Tasks at school level

Unlike at university level, here the theoretical aspects of hypothesis testing are not dealt with. Here an important issue is to try to make children understand the notion of dependence, to convince them that statistics could be and should be applied in real world problems and to develop such abilities through discovery and activity approach. Our present example offers a possibility to accomplish this job. We have to transform the tasks posed to college students, since to explain why we use the logarithmic-sum function (3) would take generally much more time than available in schools. The raw material is obviously needed, thus we have to include tasks 4.1 and 4.2. Using the output of Task 4.2, we can construct succession-tables, showing which block-letters are positioned after a given block-letter in the clear-text.

The Succession Index $SI(Y|X)$ is defined as the index-number of the letter Y in the conditional frequency order of Y among the letters following X in the source sequence. E.g. $S(Y|X)=1$ iff Y is the most frequent letter to follow up the letter X, it is 2, if Y is the second most frequent letter, and so on. It is the teacher's job to explain and demonstrate the utility of the index numbers. Refer to Nemetz (1991) for an explanation of this. The matrix S with entries $SI(Y|X)$ is called the succession matrix. The statistics (3) used for making the decision should be replaced accordingly by

$$[S(C(1,J) \,/\, C(1,K)) + S(C(2,J) \,/\, C(2,K)) + \ldots + S(C(N,J) \,/\, C(N,K))] \,/\, N. \qquad (4)$$

The decision algorithm remains the same with this exception.

Now we formally list the tasks in accordance with section 4.

Task 5.1 is identical to Task 4.1.

Task 5.2 is the same as Task 4.2., with the possible restriction to the 1-step transitions.

Task 5.3. Prepare the succession matrix *S* of Succession Indices based on the output frequency table(s) of the Task 5.3.

Task 5.4. Give arguments/run a discussion in the classroom with the aim to show that the goodness-characteristics of the decision procedure do not depend on the choice of the permutation *P*. Conclude that therefore any performance analysis can be carried out using the identity permutation, only.

Task 5.5. Take a sample ($U(i)$, $V(i)$) of pairs of characters from the "control" text, which are separated by exactly $k - 1$ letters, for $k = 1, \ldots, 5$. Determine the empirical distribution of the statistic $SI(V | U)$. (In this case it ranges from 1 to the alphabetsize through integers). Find a graphical representation that can convincingly show, why the statistics (4) works.

Task 5.6. Same as Task 4.6., with the difference in the statistics used.

Bibliography

Hawkins A. (1989). The annual United Kingdom Statistical Prize, in M. Morris (ed.), *The teaching of statistics, Studies in mathematics education*, Vol. 7, Unesco.

Kaeding F. W. (1898). *Haufigkeitswoerterbuch der deutschen Sprache*, Akademic-Verlag, Berlin.

Kahn D. (1967). *The codebreakers*, MacMillan Co., New York.

Nemetz T. (1972). *Information type measures and finite decision problems*, Lecture Notes, No. 2/1972, Carleton University, Ottawa.

Nemetz T. (1991), Automatic Discrimination of Written Languages from Random Strings with an Application to Error Detection, in M. Niss (ed.), *Teaching Mathematical Modelling and Applications*, Ellis Horwood, New York, pp. 230-241.

Nemetz T. (1993), Cryptology: a Rich Source of Applications Offering Entertaining Mathematical Instruction, in J. de Lange *et al.* (eds.), *Innovation in Maths Education by Modelling and Applications*, Ellis Horwood, New York, pp. 93-102.

Appendix

An illustration to tranpositional secret codes. This example provided our students with valuable experience in a statistics course for student teachers at METU. Should you decide to use it in your own course, we would be glad to share your experience. The following text was distributed:

```
We have used capital letters TO WRITE AN ENGLISH TEXT. THEN ALL SPACES
HAVE BEEN REPLACED BY THE LETTER X.
ANDXTHEXRE  SULTINGXTE  XTXHASXBEE  NXDEVIDEDX  INTOXBLOCK  SXOFXLENGT
HXTENXWHIL  EXALLXPUNC  TUATIONXMA  RKSXHAVEXB  EENXOMITTE  DXTHENXWEX
HAVEXAPPLI  EDXTHEXSAM  EXPERMUTAT  IONXTOXALL  XBLOCKSXBE  LOWXWEXGIV
EXTHEXRESU  LTXYOUXARE  XINVITEDXT  OXRECONSTR  UCTXTHEXOR  IGINALXTEX
TXTHEXFIRS  TXCORRECTX  SOLUTIONXI  SXWORTHXOF  XFIVEXPOIN  TSXTIESXWI
LLXBEXDISS  OLVEDXBYXR  ANDOMXCHOI  CEXINDEPEN  DENTXESSAY  SXLEADINGX
TOXGENERAL XSOLUTIONX METHODSXWI LLXBEXHONO UREDXBYXTE   NXPOINTS
**************** an encoded message follows ****************
OJTCHTXEBE  XOTXVISRFE  AISUEHVOBX  ERXHRXUCAE  VTSECAIIIT  XENHAXDCTN
XEIVCILRSA  CNCUECXDOS  BXEHETXTYD  VRTIUXIGEN  NBUXXYAEXC  RZXDMUAFIM
OLSWSAFOLX  CNIRTXXTOO  XOHTUBEXTT  VNMEXEDCAA  OXICNEXSPT  XNDXNEFDAC
OMTNVELEPE  TCONOXXHEF  XOOCOLYXTG  BTTXTNIEUR  TOLAXOANIN  EOEMDXVFLE
YMIBTNBOXX  NXDNILIAGZ  ETEVEXFICF  SNTXYLUGIX  EOCREHRUSX  FTXESEOHXX
ESYTNUVIRI  MNUPTXXIAO  XPILALEPAT  RECRDERASX  XXAEXHUMAA  FSMISNOTXX
IGNALUTXNA  SINIXDNRPI  SCERXGAXIB  CXXOESROHA  AEHTRPPXRE  CSRAXEESEN
CGUOXYARKB  OXTNDNFIRX  SIILREIPCD  XPORANYPAR  SAXDCAENXH  EAOIOCPTRO
XEAEXNNSRI  ADOTCRXXNH  DTXELXAHXE  LXXFERUOTS  ACTXEREHRS  ESGAXOHTXT
XPILXEFFAO OXXXACIXNT
```

BEST COPY AVAILABLE

PROPORTIONS, PROBABILITY AND OTHER MATTERS

Flavia Jolliffe
School of Mathematics, Statistics and Computing
University of Greenwich, Woolwich, London SE18 6PF, UK

1. Introduction

A study to investigate students' facility with proportions has been undertaken by the author and Fay Sharples of the University of Waikato in New Zealand over the period 1989 to 1992. The initial study was done during 1989 and 1990 and 64 students at the University of Waikato and 57 at Brunel University in the UK took part. Some results of this study have been reported elsewhere (Jolliffe and Sharples, 1991; Jolliffe and Sharples, 1993).

Students were given a short set of self-completion questions to test their understanding of proportion. They were instructed to "Give the answers which you *think* and *feel* are correct" and to write DK if they felt they did not know the answer. In all questions they were also asked to give a brief reason why they had chosen the answer they gave. The test took students about 25 minutes to complete.

All students in the initial study were at the start of their university courses at the time they were tested. Apart from 26 Brunel students who were registered for degrees in a mathematical subject, the students' main subject of study was in another discipline, but they were required to study some statistics. Many of these students were weak in mathematics whereas others had a good foundation in school mathematics and some had had instruction in probability. In spite of the variety of mathematical backgrounds and abilities the spread of responses was broadly similar in the two institutions and responses of the more mathematical students were not noticeably different from those of the others.

There was a tendency for students to think that events were equally likely, a phenomenon noticed also by others (e.g. Lecoutre, 1993), but more surprising and worrying was the relatively large number of incorrect responses due to arithmetical errors, to careless reading of the questions, and to an apparent lack of understanding of terms such as *estimate* and *product*.

2. A follow-up

We made some changes to our questionnaire after studying the results of the initial study. We made it slightly shorter, cutting out some questions which we felt were unsatisfactory in some way, such as two new versions of the "two children" problem where the chance that the other child is the opposite sex from the one observed is invariably said to be 1/2 regardless of the situation under consideration (Bar-Hillel and Falk, 1982).We amended the wording of some questions in the hope that they would then be clearer and we added a question asking the students the highest level (in terms of public examinations) at which they had studied mathematics previously. We kept the same instructions as before, and, as in the first version of the test, all questions were followed with "Why do you say this?" and nearly all were open-ended.

In the Spring of 1992, 127 students in New Zealand and 29 students in the UK, all of whom were taking statistics as a service course, completed the revised version of the questionnaire. As previously the New Zealand students were at the University of Waikato. Just over 2/3 of them had studied for a mathematics Bursary examination (taken in the final year at school) but would not have achieved a high mark in it, and about 70% were first year students. A random sample of 30 questionnaires completed by students at Waikato was taken for the purpose of this paper.

There were several differences between the UK students in the 1992 study and those in the earlier study. They were at a different institution – Thames Polytechnic – which has since become the University of Greenwich; reading for a different degree – Environmental Health – which involved them in acquiring expertise across a range of subjects, for example, Physics, Statistics, Microbiology, Technology, Law, Occupational Health and Safety; they were halfway through their first year although at the start of the Statistics course, and several of them were mature students with some work experience. Very few of them had studied mathematics beyond the level usually examined at age 16. The UK polytechnics traditionally taught more vocational courses than the universities and gave places to students having a wide range of abilities and entrance qualifications. In practice this meant that the more able students academically tended to go to universities as a polytechnic degree was thought in many circles to be second best. It is too early yet to say whether granting polytechnics university status in the Summer of 1992 has had any marked effect on the entrance standard.

As with the previous study, the results in this follow-up were interesting and not always what we expected. This paper discusses the results on three questions.

3. Numbered balls

The second question was about a blindfolded person taking a ball from a bag containing 10 balls which are identical except for their labels which are 9, 3, 6, 2, 7, 5, 0, 4, 8, and 1 (listed in the same order as in the question - see Appendix question 1) that is, each digit appears on exactly one ball.

Some students apparently did not understand the significance of *or* in this question and 10 thought there were two separate answers to part (a), 8 that there were two answers to part (b). The use of "or" in probability theory with its interpretation of "at least one of" is not a natural usage. When we drafted this question we were concerned with students' understanding of "at least" and "less than" and to some extent with how they would treat 5 which is included in the event "less than 7". We had expected them to count up how many numbers satisfied the criteria and divide this by 10, but many of those who gave a single answer first found two separate probabilities. They did not always then combine them correctly, for example two students multiplied them instead of adding.

Sixteen of the Waikato students in the sample but only 9 of the Thames students obtained the correct answer to part (a). The corresponding figures for (b) were 16 and 10. As expected there were some misunderstandings of "at least 8". Four of the Waikato students in the sample and two of the Thames students appeared to interpret it as "8 or less", that is, "at most 8", giving an answer of 9/10 to part (a). There were also other interpretations of the question such as "exactly 8" and "greater than 8". One Thames student said there is a probability of 2 in 9 of obtaining a number which is at least 8, and another guessed that the probability was 4 in 10. Three of the 59 students were unable to give an answer.

In part (b) four students clearly thought that "less than 7" meant "7 or less" as they stated that there were 8 out of 10 chances of getting a number less than 7. If they correctly realised that 5 was included in the event their final answer was 8/10. Another four students added the probabilities of 1/10 for a 5 and 7/10 for a number less than 7, that is, double counted the 5, and thus they too obtained a final answer of 8/10. The "Why do you say this?" for part (b) enabled us to sort out the two kinds of errors here.

Eight Waikato and four Thames students appeared to think there were six numbers less than 7, presumably forgetting to count the 0. "Less than 7" was interpreted as "more than 7" by a few and a student at Thames justified the answer 2/10 with the statement (quoted as written) "Between 5 and 7. There is no balls labelled except the ball of No. 6. The ball has to be less than 7 therefore there are only 2 balls i.e. 5 and 6".

4. A question about dice

This question (see Appendix question 2) was a revised version of a question asked in the previous study where the second sentence had started "The *product* of the two numbers". It was clear from responses in the earlier study (Jolliffe and Sharples, 1993) that many students thought that a product was a sum or did not know what was meant by a product. The changed wording appeared to have removed this confusion as only 2 students in the 59 gave any indication of considering sums and only two students felt unable to answer. However, as in the previous study, the modal response was "Equally likely" selected by 32 of the 59 students (54%) and 20 of these said that even and odd numbers on a single die are equally likely in explaining their responses. The response "is more likely to be odd than even" was selected by 5 students. The students at Thames were more inclined to choose "Equally likely" and less likely to choose the correct result of "Even" than the Waikato students. A summary of the results is given in Table 1 with combined figures for the previous study for comparison. Had students been aware that any number multiplied by an even number is even this would have been a very easy question. Would it have made any difference if we had lead them into this?

5. Sums and discs

This question (see Appendix question 3) was also asked in the previous study, but for the follow-up we put the player with two winning totals first as we had wondered whether some students in the previous study had selected the player with three winning totals as the one more likely to win because this was the first listed. As with the dice question, Waikato students did better than Thames students on this question. Thirteen of the sample of 30 Waikato students correctly thought that A had more chance of winning whereas only 5 of the 29 Thames students thought that A had more chance - 31% correct overall compared with 36% correct in the previous study. Eleven of the 13 and 2 of the 5 gave a correct argument for choosing A. Six of the Waikato students and 9 of the Thames students chose equally likely, and 9 of the Waikato students and 13 of the Thames chose B. In explaining their (wrong) choice of B as the more likely winner 12 students said that B has more winning totals, and 6 thought there were only six ways of combining the numbers on the discs that is, they did not realise there are two ways of obtaining the same total when the numbers on the two discs are different.

6. Conclusions

Our first impression of the results of this study is that there were more wrong and strange answers than in the previous study, and possibly a higher proportion of no responses and "Don't knows" in some questions. There are a number of possible explanations for this - changes in the questions asked, known differences between the groups in the two studies, and a possible cohort effect reflecting changes in syllabuses and standards between 1989 and 1992. Although changes in the questions are in the control of the experimenter, it is difficult to collect information about and make adequate allowance for confounding factors in a study of this nature where one is relying on the willingness of the students to answer the questions. As is well known, it is difficult to experiment in education and much of the research in statistical education has taken place with a small number of volunteers in an artificial setting (see Hawkins *et al.*, 1992).

This does not mean that research is worthless. A study such as ours is more in the nature of an exploratory case study. It attempts to investigate the extent of common misunderstandings and to probe into reasons for them. Students tend to dislike probability questions and become disheartened when their answers are wrong. By finding out more about why those answers are wrong helps us to improve the situation by, for example, making sure that students understand both the language of probability theory questions and the mathematical terms used in them.

Bibliography

Bar-Hillel M. and Falk R. (1982), Some teasers concerning conditional probability, *Cognition*, 11, pp. 109-122.

Hawkins A., Jolliffe F. and Glickman L. (1992), *Teaching statistical concepts*, Longman, London.

Jolliffe F. and Sharples F. (1991), An investigation of the knowledge of proportion and probability that students bring to university with them, abstract in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol.1, p. 370.

Jolliffe F. and Sharples F. (1993), Proportions and products - what do students know about these? *Teaching Statistics*, 15(3), pp. 78-80.

Lecoutre M-P. (1993), Cognitive models and problem spaces in "purely random" situations, *Educational Studies in Mathematics*, 24, pp. 559-568.

Appendix - Questions discussed in this paper

Note: Each question was followed by "Why do you say this?"

1. A ball is to be drawn by a blind-folded person from a bag containing ten balls. The balls are identical except for their labels which are:

9, 3, 6, 2, 7, 5, 0, 4, 8 and 1.

(a) What is the chance that the ball chosen is labelled 5 *or* with a number which is *at least* 8?

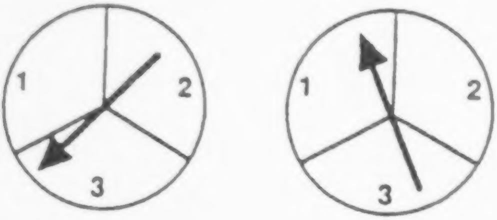
(b) What is the chance that the ball chosen is labelled 5 *or* with a number *less* than 7?

2. Two ordinary 6-sided dice are rolled. If the two numbers are *multiplied* together the answer is:

more likely to be odd than even ☐
more likely to be even that odd ☐
as likely to be even as odd ☐

Tick the box by the answer you select

3. Two discs are marked with numbers.



Each disc has a pointer which spins round. A game is played by spinning both pointers and then the numbers where they stop are *added*.

Player A wins if the *total* is 4 or 5.
Player B wins if the *total* is 2 or 3 or 6.

Which player has the greater chance of winning, or do they both have the same chance of winning?

Table 1 - *Results of the dice question*

	1992				1989/90	
	Waikato		Thames			
	N.	%	N.	%	N.	%
Odd	3	10	2	7	6	5
Even*	13	43	7	24	40	33
Equally likely	14	47	18	62	64	53
DK or no response	0	-	2	7	11	9
Total	30	100	29	100	121	100

* Correct answer

TEACHING THE ESSENTIALS OF ELEMENTARY PROBABILITY THEORY OR HOW NOT TO CALCULATE THE MEAN OF THE NEGATIVE HYPERGEOMETRIC DISTRIBUTION

Peter Nüesch
Federal Institute of Technology, Mathematics Department
MA-Ecublens, CH-1015 Lausanne, Switzerland

1. Introduction

Brosch (1991) considers a binary urn model with r red balls out of a total of N. He lets his students simulate the length X of the game if one draws, without replacement, till k red balls have appeared. The students came up with the correct answer

$$E(X) = k \cdot \frac{N+1}{r+1}.$$

He obviously does not attempt to obtain this expectation in class but invites readers to do so. The two submitted solutions (Bach, 1991; Janous, 1991) use the probability law of X and calculate the mathematical expectation by direct summation of $\sum x\, P(X = x)$.

In the following we present a solution which reduces the problem to a simpler one, relates it to another problem that uses combinatorics sparingly and which has been assigned to the 15 to 18 year age group at an International Mathematics Olympiad. Not only does this make it accessible to students of a first elementary course in probability theory but at the same it gets across to the student a feeling of the randomness of the problem. This, of course, is a basic principle of any course in probability. A mere simulation as in (Brosch, 1991), or sophisticated summations as in (Bach, 1991) and (Janous, 1991), however beautiful mathematically, simply do not convey any sense of randomness. Along the way, many of the basic notions of elementary probability are used. Our solution then can be presented as a review example of these basic notions or as an accompanying example while introducing them.

In the following we will adapt the notation used in Bach (1991), Brosch (1991) and Janous (1991) to the accepted international standards as in Feller

(1968). In particular, *X* will denote the random variable and not the number of red balls in the urn.

2. Reducing the problem to a simpler one

It is clear that the waiting time *X* till the occurrence of the *kth* success is the sum

$$X = X_1 + X_2 + \ldots + X_k, \tag{1}$$

where X_i is the waiting time after the $(i-1)st$ till the *ith* success. The proportionality of $E(X)$ with respect to *k* lets one hypothesize that the X_i are identically distributed. While a student has no difficulty in accepting this fact when sampling is done with replacement, he does have difficulty in exhaustive sampling. In an elementary course we would not try to prove the identical distribution property. The following easy example (Ross, 1988, p. 241, problem 2 and p. 323, problem 3) shows beautifully that it is so.

A bin of 5 transistors is known to contain two that are defective. The transistors are to be tested, one at a time, until the defective ones are identified. Denote by X_1 the number of tests made until the first defective is spotted and by X_2 the number of additional tests until the second defective is spotted; i) find the joint probability mass function of X_1 and X_2; ii) find the expected number of tests that are made.

While i) is helpful to establish dependence between the X_i, it is really ii) that interests us.

The sample space Ω consists of the $\binom{5}{2} = 10$ arrangements of the 5 transistors (*D* for defective, *N* for non-defective).

					x_1	x_2	y_1	y_2	y_3
D	*D*	*N*	*N*	*N*	1	1	3	1	1
D	*N*	*D*	*N*	*N*	1	2	2	2	1
D	*N*	*N*	*D*	*N*	1	3	2	1	2
D	*N*	*N*	*N*	*D*	1	4	2	1	1
N	*D*	*D*	*N*	*N*	2	1	1	3	1
N	*D*	*N*	*D*	*N*	2	2	1	2	2
N	*D*	*N*	*N*	*D*	2	3	1	2	1
N	*N*	*D*	*D*	*N*	3	1	1	1	3
N	*N*	*D*	*N*	*D*	3	2	1	1	2
N	*N*	*N*	*D*	*D*	4	1	1	1	1

Table 1

The probability laws for X_1 and X_2 are identical. (We observe the same fact for the random variables Y_j, the respective waiting times till the non-defective transistors show up). Here is an example of two (three) different random variables defined on the same sample space having the same distribution.

3. The probability law of X

First we restrict our attention to X_1. In order to simplify notation let us omit for the moment the subscript 1.

The event $\{X > x\}$ means that the waiting time up to success 1 is larger than x, or, equivalently, that the first x draws are failures. Thus the tail probability $P(X > x)$ is

$$P(X > x) = \frac{N-r}{N} \cdot \frac{N-r-1}{N-1} \cdot \frac{N-r-2}{N-2} \cdots \frac{N-r-x+1}{N-x+1}$$

$$= \frac{(N-r)_x}{(N)_x} . \tag{2}$$

Observe the analogy to the geometric case, i.e. when sampling is done with replacement

$$P(X > x) = \frac{(N-r)^x}{N^x} .$$

It follows that

$$P(X = x) = P(X \le x) - P(X \le x-1)$$

$$= 1 - P(X > x) - 1 + P(X > x-1)$$

$$= P(X > x-1) - P(X > x)$$

$$= \frac{(N-r)_{x-1}}{(N)_{x-1}} - \frac{(N-r)_x}{(N)_x}$$

$$= \frac{(N-r)_{x-1}}{(N)_x} r \quad , \qquad x = 1, \ldots, N-r+1. \tag{3}$$

Again, observe the analogy to the geometric distribution

$$P(X = x) = \frac{(N-r)^{x-1}}{N^x}\, r = q^{x-1} p \quad x = 1,\ 2, \ldots$$
$$\text{where } p = \frac{r}{N} \text{ and } q = 1 - p.$$

(Most elementary texts treat the problem of the N identical keys of which only one opens a door. To the student's surprise the probability that the door opens on any one attempt is $1/N$ if keys already used are discarded. This is formula (3) for $r = 1$).

The case $k > 1$ leads to the negative binomial distribution when sampling is done with replacement

$$P(X = x) = \binom{x-1}{k-1} \frac{r^k (N-r)^{x-k}}{N^x}, \quad x = k,\ k+1, \ldots$$

and for sampling without replacement one obtains the analogous law

$$P(X = x) = \binom{x-1}{k-1} \frac{(r)_k (N-r)_{x-k}}{(N)_x}, \quad x = k,\ k+1, \ldots,\ N-r+k. \qquad (4)$$

It is the expectation of this distribution which we are asked to find. In the literature, e.g. Johnson and Kotz (1977), it has received the name "negative hypergeometric" distribution, in analogy to the fixed sample distributions with a variable number of successes, where "binomial" is used if sampling is done with replacement, and "hypergeometric" for sampling without replacement. (Strickly speaking, the name negative binomial was given to the number of failures before success k occurs. This stems from the fact that the binomial coefficient can be written as $\binom{-r}{k}$). Patil and Joshi (1968) make a further distinction between a negative and an inverse hypergeometric distribution).

We will not need the law for $k > 1$. If we can find the mean of X_1 and thus of X_i, *the mean of* $X = \sum^k X_i$ *can be obtained without explicitly knowing the probability law of* X.

4. Relation to another problem

One finds in Ross (1988, p. 145, exercise 20) the following problem:

Balls numbered 1 through N are in an urn. Suppose that r, $r \leq N$, of them are randomly selected without replacement. Let Y denote the largest number selected. Find the probability mass function of Y.

To obtain the probability mass function of Y note that if y is the maximum number drawn all $r-1$ other balls in the sample must carry numbers less than r. Thus

$$P(Y=y)=\frac{\binom{y-1}{r-1}}{\binom{N}{r}}, \quad y=r, \ldots, N.$$

Since the probabilities add up to 1 we get the combinatorial identity

$$\sum_{y=r}^{N}\binom{y-1}{r-1}=\binom{N}{r} \tag{5}$$

which is the famous sum of entries on a 45° diagonal of the Pascal triangle.

We can equally easily compute the probability mass function of Z, the smallest number selected. By the same argument used for the law of Y, $r-1$ numbers are to be selected from the $N-z$ numbers larger than z. Thus

$$P(Z=z)=\frac{\binom{N-z}{r-1}}{\binom{N}{r}}, \quad z=1, \ldots, N-r+1.$$

The probability space Ω for $N=5$ and $r=2$ is

		y	z	$y-z$
1	2	2	1	1
1	3	3	1	2
1	4	4	1	3
1	5	5	1	4
2	3	3	2	1
2	4	4	2	2
2	5	5	2	3
3	4	4	3	1
3	5	5	3	2
4	5	5	4	1

Table 2

By comparing Table 2 to Table 1 one observes that

$$P(X_1 = x) = P(Z = z)$$

and

$$P(X_2 = x) = P(Y - Z = t).$$

The variable X_1, the position of the first detective item of r defective ones in a total of N, has the same probability law as the minimum Z of all samples of size r drawn without replacement from a total of N numbered items. Of course, this result can be obtained in a purely formal way. Instead of transforming (3)

$$P(X = x) = \frac{(N-r)_{x-1}}{(N)_x} \cdot r$$

into

$$P(X = x) = \frac{\binom{N-x}{r-1}}{\binom{N}{r}}, \quad x = 1, 2, \ldots, N-r+1 \tag{6}$$

it is easier to work with tail probabilities. We have (2)

$$P(X > x) = \frac{(N-r)_x}{(N)_x}$$

$$= \frac{\binom{N-r}{x}}{\binom{N}{x}} \qquad \text{using } (a)_b = \binom{a}{b} b! \tag{7}$$

$$= \frac{\binom{N-x}{N-r-x}}{\binom{N}{N-r}} \qquad \text{using } \binom{n}{j}\binom{j}{i} = \binom{n}{i}\binom{n-i}{j-i}$$

$$\text{for } n = N, i = x, j = N-r$$

$$= \frac{\binom{N-x}{r}}{\binom{N}{r}}, \quad x = 1, 2, \ldots, N-r \tag{8}$$

and

$$P(X = x) = P(X > x-1) - P(X > x)$$

$$= \frac{\binom{N-x+1}{r} - \binom{N-x}{r}}{\binom{N}{r}}$$

$$= \frac{\binom{N-x}{r-1}}{\binom{N}{r}}$$ using the basic relation of the Pascal triangle.

Remark: If one proceeds this way for the negative hypergeometric probability law (4), one obtains

$$P(X = x) = \frac{\binom{x-1}{k-1}\binom{N-x}{r-k}}{\binom{N}{r}} . \tag{9}$$

This reduces for $k = 1$ to the form (6) obtained above.

5. The mathematical expectation of X

We take all samples of size r and order the numbers

$$x_1 < x_2 < \ldots < x_h < \ldots < x_r .$$

One has

$$P(X_h = x) = \frac{\binom{x-1}{h-1}\binom{N-x}{r-h}}{\binom{N}{r}} , \quad x = 1, \ldots, r \tag{10}$$

since h – 1 numbers are smaller than x and $r - h$ numbers are larger than x. Since this is a probability law, and the probabilities are therefore summing up to 1, we found the combinatorial relation

$$\sum_{x}\binom{x-1}{h-1}\binom{N-x}{r-h}=\binom{N}{r}. \tag{11}$$

Observe that this relation is the sum of the negative hypergeometric probabilities (9) for $h = k$. In the literature, (11) is called the Vandermonde convolution, e.g. Riordan (1968, p. 8).

By using (11) and

$$\binom{a}{b}=\frac{a}{b}\binom{a-1}{b-1} \tag{12}$$

we get

$$\begin{aligned}E(X_1)&=\frac{\sum x\binom{N-x}{r-1}}{\binom{N}{r}}\\&=\frac{N+1}{r+1}.\end{aligned} \tag{13}$$

(13) holds for all i, i.e. $E(X_i) = (N+1)/(r+1)$. For the random variable X of (1) we get therefore

$$E(X)=\sum_{i=1}^{k}E(X_i)=k\frac{N+1}{r+1}.$$

Bibliography

Bach G. (1991), Lösung 2 zu "Eine Formel für den Erwartungswert", *Praxis der Mathematik*, 33, pp. 279-281.

Brosčh H. (1991), Eine Formel für den Erwartungswert, *Praxis der Mathematik*, 33, pp. 132-133.

Feller W. (1968), *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd edition, Wiley, New York.

Janous W. (1991), Lösung 1 zu "Eine Formel für den Erwartungswert", *Praxis der Mathematik*, 33, pp. 278-279.

Johnson N.L. and Kotz S. (1977), *Urn Models and their Applications*, Wiley, New York.

Patil G.G. and Joshi S.W. (1968), *A Dictionary and Bibliography for Discrete Distributions*, Oliver & Boyd, Edimburgh.

Riordan J. (1968), *Combinatorial Identities*, Wiley, New York.

Ross S. (1988), A *First Course in Probability*. 3rd edition, Macmillan, New York.

STATISTICAL SIMULATIONS

Mary Rouncefield
Chester College of Higher Education
Cheyney Road, Chester, CH1 4BJ, UK

1. What is a statistical simulation?

A statistical simulation is a model of a real situation in which chance or random factors are involved. In a queuing situation, for example, a simulation can be used to help answer questions such as:

- How many pumps are needed at a petrol station?
- How many pay booths are required at a toll bridge?
- Is a central queuing system more efficient than queues at separate counters in a bank or post office?
- Will the introduction of an appointments system mean that a doctor's patients will have to wait shorter lengths of time for consultation?

In order to answer these questions the simulation will need to model both the arrivals of cars, or customers or patients, and the time taken to "process" each one. Chance elements come into play at both ends of the queue, arrivals and departures, and these chance factors must be built into the model. This is a fairly advanced statistical simulation and is explained in Rouncefield and Holmes (1989).

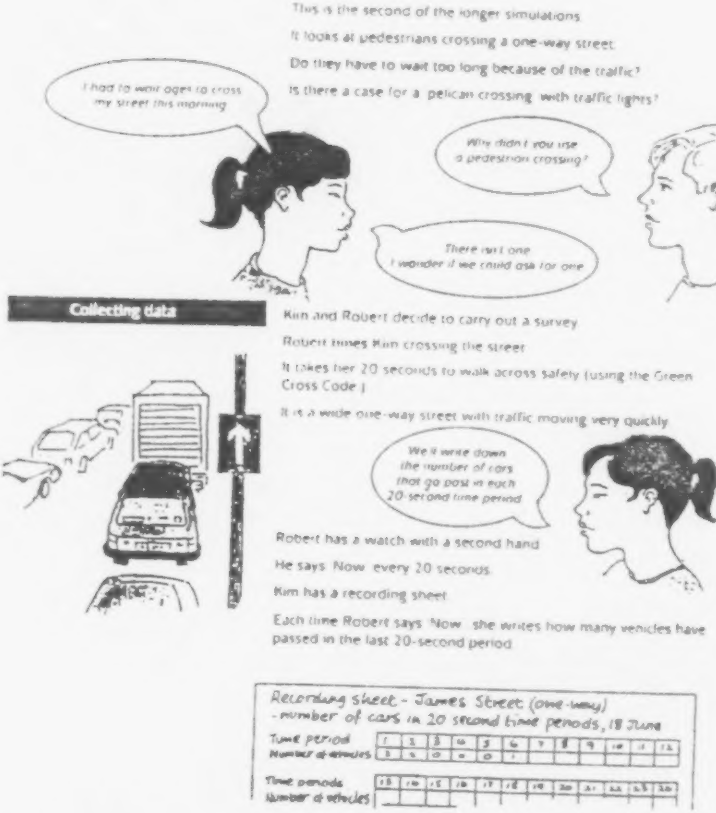
One way street simulation

The simplest traffic simulation can be based on a "one way" street. Here the problem is described in Fig. 1.

Pupils can use the data provided in the text or they can collect their own data on a one way street near their school. First of all, the results can be examined to find the longest time a pedestrian would have to wait to cross the road. In order to get a fuller picture, more data is needed, and this will take a considerable time to collect, we can generate further results in a simulation. In order to do this a frequency table can be drawn up and relative frequencies calculated. These relative frequencies can be used as probabilities and further results generated using random numbers or a spinner as in Fig. 2.

The class can then discuss the original problem: "Is a pedestrian crossing needed"? A simulation can then be run to see what would happen if a pedestrian crossing was introduced:

- How long will pedestrians have to wait? (on average).
- How long will vehicles have to wait? (on average).


I had to wait ages to cross my street this morning
Why didn't you use a pedestrian crossing?
There isn't one I wonder if we could ask for one
We'll write down the number of cars that go past in each 20-second time period

This is the second of the longer simulations

It looks at pedestrians crossing a one-way street

Do they have to wait too long because of the traffic?

Is there a case for a pelican crossing with traffic lights?

Collecting data

Kim and Robert decide to carry out a survey

Robert times Kim crossing the street

It takes her 20 seconds to walk across safely (using the Green Cross Code.)

It is a wide one-way street with traffic moving very quickly

Robert has a watch with a second hand

He says Now every 20 seconds

Kim has a recording sheet

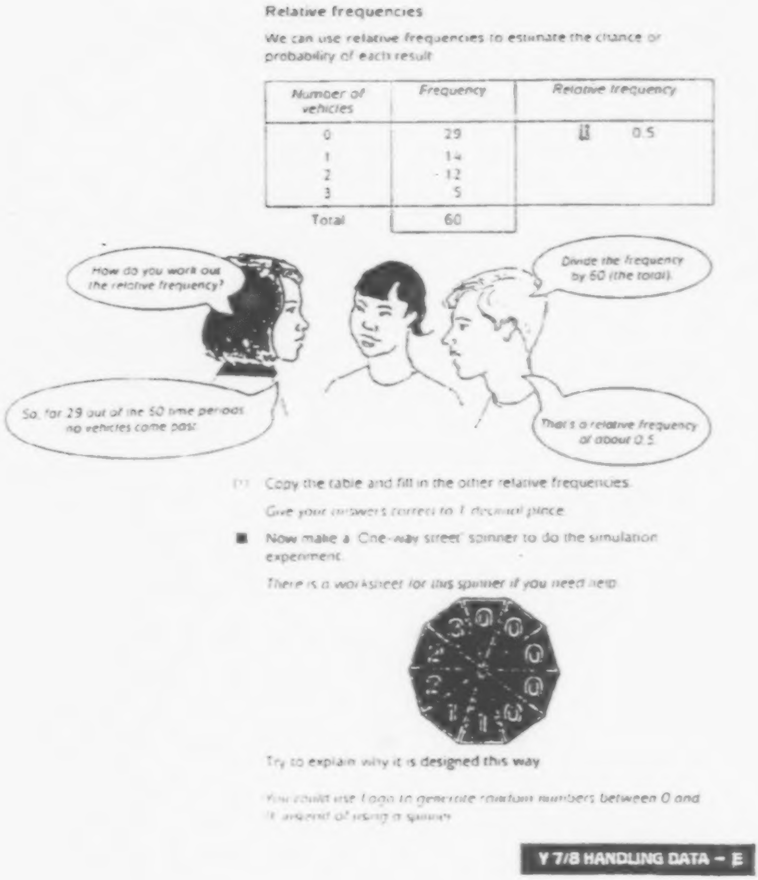
Each time Robert says Now she writes how many vehicles have passed in the last 20-second period

Recording sheet - James Street (one-way)
- number of cars in 20 second time periods, 18 June

Time period	1	2	3	4	5	6	7	8	9	10	11	12
Number of vehicles	3	[illegible]	0	[illegible]	0	1						
Time periods	13	14	15	16	17	18	19	20	21	22	23	24
Number of vehicles												

Figure 1. *One-way street*

BEST COPY AVAILABLE



Relative frequencies

We can use relative frequencies to estimate the chance or probability of each result

Number of vehicles	*Frequency*	*Relative frequency*
0	29	$\frac{29}{60}$ 0.5
1	14	
2	12	
3	5	
Total	60	

How do you work out the relative frequency?

So, for 29 out of the 60 time periods no vehicles come past.

Divide the frequency by 60 (the total).

That's a relative frequency of about 0.5.

(1) Copy the table and fill in the other relative frequencies.

Give your answers correct to 1 decimal place.

■ Now make a 'One-way street' spinner to do the simulation experiment.

There is a worksheet for this spinner if you need help.

Try to explain why it is designed this way

You could use Logo to generate random numbers between 0 and 9 instead of using a spinner.

Y 7/8 HANDLING DATA – E

Figure 2. *Relative frequencies*

Simulation and reality.

This "One Way Street" simulation has some important features:

1. Real data are used.
2. Relative frequency is used as an estimate for probability.
3. Statistical and probability ideas and skills are linked and used in conjunction. In many cases, pupils work either with probability or

statistics, rarely with both. In the minds of many pupils (and teachers) statistics and probability are separate entities and links between them rarely made. (One exception being the tabulation and analysis of a probability experiment).

Two way street simulation

Using the skills and ideas acquired in the One Way Street simulation, pupils can progress to a more complex simulation; a two way street in which:

- Traffic approaches from two directions. Two sets of observations are required to generate probabilities so that a two-way flow can be simulated.
- Pedestrians arrive at both sides of the street, wishing to cross over. Again two sets of data are needed to simulate this two-way flow.

Altogether 4 flows are involved in the simulation. A team of pupils can work together using a map and toy cars and figures as shown in Fig. 3.

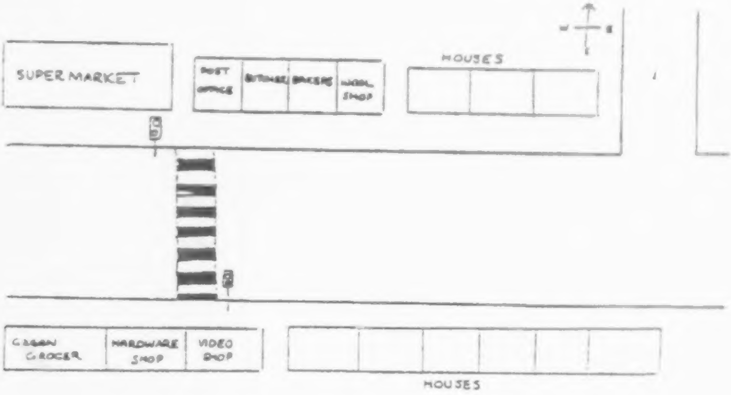


Figure 3. *Two way street*

The simulation can be run both without and then with the pedestrian crossing.

A more enjoyable, although possibly more chaotic version of this simulation, can be run by drawing a roadway on the playground in chalk. Pupils themselves can take the roles of pedestrians and cars.

2. Population models

Many pupils are interested in ecology and conservation issues. They can use library books to research data on the breeding and mortality of endangered species. As ornithologists are such enthusiastic observers and collectors of information, some of the best data available relate to bird populations.

Fig. 4 shows a flow chart showing the processes involved in modelling a 2 year breeding cycle for a pair of Royal albatross.

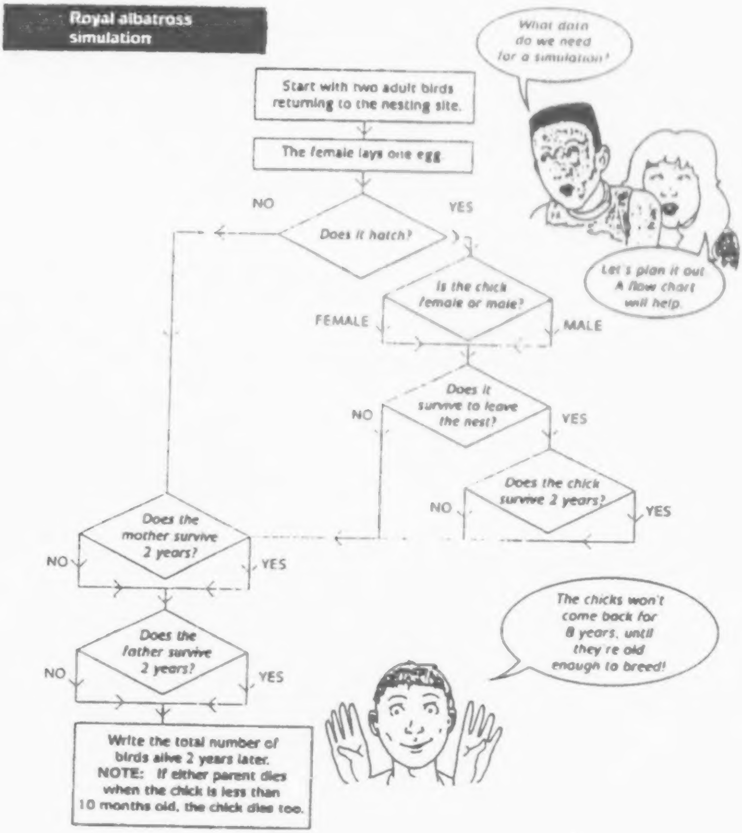

Royal albatross simulation
What data do we need for a simulation?
Start with two adult birds returning to the nesting site.
The female lays one egg
NO
YES
Does it hatch?
Let's plan it out. A flow chart will help.
Is the chick female or male?
FEMALE
MALE
Does it survive to leave the nest?
NO
YES
Does the chick survive 2 years?
NO
YES
Does the mother survive 2 years?
NO
YES
The chicks won't come back for 8 years, until they're old enough to breed!
Does the father survive 2 years?
NO
YES
Write the total number of birds alive 2 years later.
NOTE: If either parent dies when the chick is less than 10 months old, the chick dies too.

Figure 4. *Royal albatross simulation*

BEST COPY AVAILABLE

Each decision box in the flow chart requires randomised results based on probabilities calculated from real data.

3. The use of statistical simulations in the classroom

In these statistical simulations, probability and statistics appear as interwoven strands of the same theme rather than as disjoint areas of mathematics. Real situations are studied and real data are used.

Mathematics is applied in real context of interest to the pupil. Statistical simulations provide the opportunity for cross-curricular projects. Situations can be modelled from numerous subject areas including: Ecology, Biology, Population Studies, Business Applications, Physics and Mechanics.

Students are also motivated to write simple programs for graphical calculators or micro computers.

Acknowledgements

Figures 1, 2, and 3 are taken from the *Handling Data Extension Year 7/8* book, part of the *Century Maths* materials published by Stanley Thornes. Figure 4 has been reproduced from the *Handling Data Extension Year 9* book in the same *Century Maths* series.

Bibliography

Hudson B. and Rouncefield M. (1991), *Handling Data Extension Year 7/8*, from the Century Maths series, Cheltenham, Stanley Thornes, U.K.

Hudson B. and Rouncefield M. (1992), *Handling Data Extension Year 9*, from the Century Maths Series, Cheltenham, Stanley Thornes, U.K.

Rouncefield M. and Holmes P. (1989), *Practical Statistics*, Macmillan Education, London.

TEACHING STATISTICS: MAKING IT MEMORABLE

Eric R. Sowey
Department of Econometrics University of N.S.W.
P.O. Box 1, Kensington, N.S.W. Australia 2033

1. Introduction

How often do teaching and learning statistics at university seem to be activities directed primarily to students' passing assessments and gaining paper credentials. This tends to be the case especially in service courses, where, indeed, the vast majority of statistics students are found. Once such an orientation becomes entrenched, longer term objectives may be regarded by teachers or students (or both) as unattainable.

If, however, we wish to promote enhanced quantitative thinking, analysis and practice in statistics-using professions, then an overriding goal of all teaching ought to be to stimulate learning that lasts.

In this paper I draw on over 20 years' experience in university teaching at every level, and in classes ranging in size from 5 to 500, to offer suggestions on ways of teaching statistics so that ideas will be long remembered[1]. I take it for granted that this requires a certain level of student understanding of those ideas, for understanding is essential to long-term learning.

To teach so that ideas will be long remembered I shall here call making teaching memorable.

2. What makes teaching memorable?

I take as my inspiration a striking aphorism attributed to the eminent psychologist, B.F. Skinner[2]: "Education is what remains when the facts one has learned have been forgotten".

---

[1] Academic tradition expects that I will provide independent corroborative evidence for my views, most desirably from formal experimental studies of what promotes student understanding and long-term learning in statistics. I have not found any detailed studies on this theme. Two very recent papers (Jolliffe, 1991 and Nitko and Lane, 1991) confirm that experimental research in this area is in its infancy.

[2] The form of words which I have used in the text and which I first heard many years ago turns out to be a pithier version of what, I have recently discovered, Skinner actually wrote (in the New Scientist, 31 May 1964, p. 484): "It has often been remarked that an educated

This leads me to ask: what does remain by way of statistical knowledge when the facts one has learned have been forgotten?

From conversations on education I have had over the past 20 years with former students, both soon after and long after they studied statistics or econometrics subjects with me, two points have crystallised. What remains with them, when the facts they have learned have largely been forgotten, is a sense of the *structure* of the subject and a sense of the *worthwhileness* of the subject[3]. Synthesising student comments, I interpret these two characteristics in the following way.

Structure is a reflection of the *coherence* of the subject and its presentation to students, and is often seen best in a *perspective* view. Worthwhileness is an amalgam of *intellectual excitement, resilience to challenging questioning,* and *demonstrated practical usefulness.*

So, to make statistics memorable these attributes of its structure and worthwhileness need to be consciously conveyed to students. How can recognition of each of these attributes translate into ways of teaching statistics? And given that there are strong linkages among these attributes, how can memorable teaching make the most of such linkages?

Before taking up these questions, I note that other attributes of good teaching are commonly mentioned. Such attributes as establishing rapport with students and providing clear explanations in class are, in the main, bound up with personal qualities of the teacher. Such qualities often make the *teacher* memorable. But I am focussing here on the principal factors that will make the *subject* memorable. These factors are not, by and large, personal qualities of the teacher, but rather *aspects of the teacher's own conception of and approach to the discipline.* My reading of the statistical education literature suggests that such factors are much less commonly discussed.

3. Structure

A hallmark of memorable teaching is to make students aware of structure in the details of a subject. It is structure which is the bridge between knowing things and understanding them, and it derives, as I have

---

man has probably forgotten most of the facts he acquired in school and university. Education is what survives when what has been learned has been forgotten".

[3] This empirically-founded insight is quite parallel to Ericksen's (1985, p. 30) dictum "Successful lecturing promotes the two basic conditions for learning and retention: motivation and meaning". My "structure" parallels meaning, and "worthwhileness" parallels motivation.

said, from all the elements of coherence in the subject. Structure is best grasped from a perspective view.

3.1 *Coherence*

To teach statistics coherently means to identify related elements in the discipline and to explain the elements in a way that highlights the relationships. In Sowey (1991) I have enlarged on this theme, identifying three dimensions of coherence which I call theme, pattern and knowledge coherence.

Theme coherence is a smooth traverse in explaining (a) the theoretical basis of statistical technique (going from assumptions to conclusions, and from simple cases to more complex ones), and (b) the transition from theory to practice. The key concept is *continuity* of logic within each theme.

Pattern coherence arises from the existence of common characteristics, or patterns, in the structure of statistical theory across different areas of the discipline. The key concept is underlying *unity* in disparate procedures.

Knowledge coherence comes from a view of statistics as woven into the tapestry of all human knowledge. The key concept is *integration* of statistics with its source and auxiliary disciplines (logic, probability, computing), its cognate disciplines (biometrics, psychometrics, econometrics, cliometrics, etc.) and its myriad disciplines of application in the sciences, social sciences and humanities.

Most textbooks aim at theme coherence in their exposition, but it is not always done with success, especially in the transition from theory to practice. On the latter count, Chatfield (1988) is a notable exception. By contrast, pattern coherence is highlighted almost exclusively in advanced texts. This is unfortunate, for there are ample instances that intermediate texts could point to (e.g. the communality of form in simple and multiple regression estimators, made apparent via the use of matrix notation; and the communality of structure in regression, experimental design, and variance component models within the general linear model). Knowledge coherence, intrinsic to making sense of the world from the statistician's standpoint, is hardly mentioned in the textbooks at all. To the statistics student, the mosaic of publications, each covering a particular inter-disciplinary area of statistical application, is no substitute for a single broadly integrated picture.

Because textbooks are generally unsupportive of teaching that pays regard to all aspects of coherence in statistics, the teacher needs to assume greater responsibility for the task. To this end, a capacity for lateral thinking about the scholarly literature and a broad general knowledge of the place of the discipline in the world of learning will stand the teacher in good stead.

3.2 *Perspective*

A perspective view woven into the exposition from time to time brings students at least three benefits: (a) it helps them chart their progress through the syllabus; (b) it promotes understanding of the coherence of the subject; (c) it makes clear what parts of the area under study are not currently being treated in detail. In other words, it lets students know what it is that they don't yet know.

Constructive though this sounds and is, many teachers do without regular pauses for perspective in their exposition. The effect is to make student learning rather like cartography in the days before flight D slow, tedious, disjointed and, ultimately, very imperfect!

4. Worthwhileness

Conveying a feeling for the worthwhileness of statistics is in every sense a key element in memorable teaching, for it is both a key motivator of students to learn well in class and the key to effective self-directed learning in the future.

Three principal qualities of the discipline need to be brought out so as to engender this sense of worthwhileness.

4.1 *Intellectual excitement*

David Attenborough, Stephen Jay Gould, Fred Hoyle, Julius Sumner Miller, David Suzuki: these names are readily recognised. What have these people in common? Each has the gift of talking about his discipline in an intellectually exciting way. I have asked myself how they generate that excitement and have carried over my conclusions to statistics.

Intellectual excitement about statistics, I can confirm from practice, will grow from teaching where:

(a) students see the discipline as one of central importance, but one in which not everything is yet settled. Intriguing questions that are still largely unanswered (e.g. the real-world validity of common parametric assumptions, the quality in small samples of asymptotically optimal inferential procedures, and the optimality of standard estimators when used after model selection via a data-directed specification search) should certainly be raised when the moment is apt whether or not a solution is being offered (and even in an introductory course);

(b) the teacher's enthusiasm for and commitment to the discipline is evident. The impact on students of these qualities cannot be overstated.

though by themselves they may do no more than ignite students' interest. To keep the fire burning, the teacher needs to show others why the subject merits *their* enthusiasm and commitment, as well;

(c) there are planned moments of surprise that will arouse curiosity and/or provoke reflection. Three well-proven teaching devices can serve here: to present an arresting example (especially one that relates to students' own life experiences), to challenge students to resolve a paradox, and to guide them to an unexpected discovery. Here are some instances.

Early in my introductory statistics course I want students to appreciate that, from samples drawn quite at random, statistics can provide inferences about populations that are in harmony with reality. As a playful musical metaphor (which should not be pressed too far) I present the results of a sampling experiment using Mozart's (1793) *Musical Dice Game*[4]. This demonstration, I have found, is not soon forgotten. Counterexamples to statistical propositions also make arresting examples (see, for example, Romano and Siegel, 1986), as do some of the more bizarre misuses of statistics in the daily press. Another lively idea comes from Mansfield (1989). Many further possibilities can be drawn from articles in *The American Statistician and Teaching Statistics.*

Paradoxes abound in statistics and, if aptly posed, can be counted on to stimulate lively and memorable discussion. Here is an example. A regression model containing an exact linear dependence among k regressors cannot be estimated by ordinary least squares *unless one regressor is deleted.* However, by transforming to (k – 1) principal components and then unscrambling the OLS regression on principal components, one has an estimated model that contains (surprise!) all k original regressors. Where is the catch? More subtle paradoxes are found in Szekely (1986).

Unexpected discoveries *that students make themselves* have the strongest impact on learning and remembering. The remarkable meaning and generality of the Central Limit Theorem, for instance, are rarely so fully appreciated as by students who have performed computer sampling simulations summarised in real-time graphics, with samples of ever-increasing size drawn from a variety of symmetric and asymmetric (even bimodal!) populations.

A similar impact is found when students investigate the geometry of the standard normal curve. They are generally astonished to discover that, if the curve were to be drawn so that the ordinate at $z = 6$ is 1 mm high, then the ordinate at the mode would be 65.7 km (yes, km) high.

---

[4] This musical curiosity presents a "treasury" of 176 detached bars of piano music. Bars are selected randomly from a table of bar numbers by sampling according to the sum of the faces of two rolled dice. Any "random composition" of 16 bars, constructed as described, is surprisingly found to be not only smoothly playable but also, indeed, quite harmoniously melodic!

4.2 *Resilience to challenging questioning*

I think of resilience under challenge as a defining characteristic of a university discipline. Displaying that resilience should, moreover, be part of the teaching of all such disciplines. Ideally, the challenging questions will come from students, themselves, but if not, the teacher ought to ask those questions.

The purpose is to elicit critique of disciplinary foundations and theoretical conventions, commentary on limitations of techniques and controversial interpretations, and scrutiny of the worth of claimed achievements. These are all, in the philosophical sense, methodological issues. The result is that a realistic three-dimensional picture of the discipline emerges.

My experience confirms that offering students (especially senior students) a methodological view of statistics in this way develops their skill in critical thinking and, by *involving them*, acts as a powerful catalyst to long-term learning.

Notwithstanding the benefits, it is, regrettably, uncommon to find support for methodological discussion in academic syllabi and textbooks. What might account for this? I have heard a number of explanations. Among them are: "it is more important to cover substantive techniques which give students a marketable skill" and "methodology is a waste of time - it never settles anything".

Such explanations too easily dismiss the subtlety of motivation to long-term learning.

The merits, for long-term learning, of a methodological approach to statistics carry over in large measure also to teaching which acknowledges the history of ideas in the discipline. By retracing just a few of the often bumpy research paths of statistical pioneers students can gain a sense of the pains and pleasures of creative thinking. What's more, they can come to appreciate the resilience, or otherwise, of the discipline to the challenging questions of an earlier generation of thinkers.

4.3 *Demonstrated practical usefulness*

Students need more than simply an assurance that statistics is important in the real world to be motivated to learn and retain statistical ideas. It is the demonstration of practical usefulness that convinces.

How can one go about such a demonstration? Illustrative practical examples, presented in the course of teaching theory, are the most obvious device. But these are hardly sufficient because they are contrived for a different primary purpose and so are usually rather unrealistic. More broadly constructive are assignments, based on real-world problems, in which students play the active role in formulating a verbally-stated problem in statistical terms, and must struggle with ill-conditioned data, decide for themselves the most appropriate technique to apply, cope with unanticipated analytical obstacles, and finally write a professional report on their investigations. This kind of problem-based learning by discovery is a powerful way to cement understanding, as the literature on using case-studies in teaching confirms. Another way to achieve some of the same outcomes with senior students is to give them a role as rostered statistical consultants to less-trained users of statistics on campus.

Most strikingly effective is to bring students into contact with statistical practitioners in government, business and industry, through one or more of these initiatives: field trips into the workplace; work-experience "sandwich" programs over several weeks; a visiting speaker seminar that is melded into the teaching program. Hahn and Schmee (1987) make an interesting proposal to take yet further the idea of a visitor from industry.

5. Linkages

I have been looking separately at five attributes of memorable teaching. Each has a place in every different kind of statistics course, but to a varying degree. For example, an *introductory service course* might emphasise theme coherence and practical usefulness, an *intermediate service course* theme and knowledge coherence, intellectual excitement and practical usefulness, an *intermediate specialist course* theme and pattern coherence and intellectual excitement, an *advanced specialist course* all the dimensions of coherence and disciplinary resilience.

Synergy, it is important to recognise, can develop from jointly emphasising these various elements. Students who respond to the discipline's intellectual excitement are, *on that account*, likely to be more motivated to seek out pattern coherence in statistical theory and to challenge disciplinary conventions. The teacher should always be alert to possibilities of fostering and forging such vital links.

All the examples I have just given have a feature in common.

Each unites one attribute from the "structure" group with at least one attribute from the "worthwhileness" group. The former group represents the wholly rational inputs into learning and remembering, while the latter also encompasses emotional influences. It is this way that linkage can make its most valuable contribution to the teaching of statistical ideas.

If it is done well, students will experience statistics as both logical and pleasurable. When their minds and feelings have both been engaged, the ideas they have learned will prove indelible.

Bibliography

Chatfield C. (1988), *Problem Solving: A Statistician's Guide*, Chapman and Hall, London.

Ericksen S.C. (1985), *The Essence of Good Teaching*, Jossey-Bass, San Francisco.

Hahn G.J. and Schmee J. (1987), Practitioner and academician co-teaching: an idea to consider, *Proceedings of the Statistical Education Section*, American Statistical Association, pp. 199-203.

Jolliffe F.R. (1991), Assessment of the understanding of statistical concepts, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 1, pp. 461-466.

Mansfield E.R. (1989), Teaching elementary probability like magi, *Proceedings of the Statistical Education Section*, American Statistical Association, pp. 87-88.

Mozart W.A. (1793), *Musikalisches Würfelspiel* [Musical Dice Game], (Köchel catalogue appendix no. 294d), Published posthumously, Schott Edition No. 4474, issued 1989.

Nitko A.J. and Lane S. (1991), Solving problems is not enough: assessing and diagnosing the ways in which students organise statistical concepts, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 1, pp. 467-47.

Romano J.P. and Siegel A.F. (1986), *Counterexamples in Probability and Statistics*, Wadsworth & Brooks, Monterey.

Sowey E.R. (1991), Teaching econometrics as though coherence matters, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 2, pp. 321-331.

Szekely G.J. (1986), *Paradoxes in Probability Theory and Mathematical Statistics*, Reidel, Dordrecht.

WHAT IS THE MAIN GOAL OF STATISTICAL EDUCATION?

Vladislav V. Shvyrkov
National University, 536 Oasis Drive
Santa Rosa, CA 95407, USA

1. Introduction

Teaching broad aspects of statistical methods with the emphasis on computational statistics is important, but it is not the main goal of statistical education. The main goal of statistical education is to teach statistical methods in accordance with the principle problem of statistics.

The principle problem of all science is to trace laws and regularities of our lives on the basis of a sample. Statistics, as a universal methodological science, participates in the solution of this problem by analysing a sample with the help of quantitative methods. Everybody knows that a statistical analysis is only as good as the data on which it is based. Therefore, the analysis of the quality of data (sample) is a principle problem of statistical science. To enable students to resolve this problem, statistical education must concentrate its attention on the critical analysis of the unwarranted (untested) assumptions (i.e. homogeneity, representativeness, and randomness) of conventional statistics.

There is an ethical aspect associated with these unwarranted assumptions in the context of statistical education. Students have the right to know when and which assumptions are unwarranted in statistical research. The neglect of such issues provides a major source of the statistical abuses perpetrated by researchers and educationists.

2. Unwarranted assumptions

Every science is built upon a few axioms, or general truths. They cannot be logically proved, but we accept them as obviously true to actual fact.

Conventional statistics is not a science since it is not built upon axioms. It is based on the following three principal unwarranted assumptions; homogeneity, representativeness, and randomness.

The first unwarranted assumption is that a sample is homogeneous. This means that we deal with good data, which is very important since our decisions are only as good as the data on which they are based. Unfortunately, statisticians do not always know what good data mean, and how that 'goodness' may be tested.

The second unwarranted assumption is that a sample is representative of the population with respect to their standard deviations. This means that the dispersion of the homogeneous sample is approximately equal to the dispersion of the population. This assumption is not, and cannot be, tested because, according to conventional statistics, a population is not known *a priori*.

The third assumption is that the expected deviations from the linear stochastic regression model (population) are a random variable comprised of unobserved residuals. This randomness assumption is attempted on *a priori* grounds because there is no formal test for unobserved residuals. The examination of the observed residuals does not necessarily help in testing the randomness assumption for the following two reasons. Firstly, the observed residuals may be affected by mis-specification of the model. Secondly, the test for normality of the observed residuals is not a test for analysing the shape of the observed distribution. It is a test which is used to detect whether or not the residuals come from the normal population.

To test the above unwarranted assumptions, statistics must be built upon four axioms (Shvyrkov,1992, 1993).

The purpose of this paper is to build a statistical science on the basis of axioms in order to test the unwarranted assumptions of conventional statistics.

3. The principal problem of statistics

The concept of the quality of a sample in statistical terms is related to the concept of the sample's representativeness with respect to a homogeneous population. How do statisticians cope with the homogeneity problem? How do they define the homogeneity of the data set? The philosophy of conventional statistics is based on the assumption that the homogeneity problem is not a statistical problem, and that "the homogeneity of the material studied cannot be ascertained in principle" (Sachs, 1982, p. 83). Therefore, in statistics, the problem of the homogeneity of the data set is treated as only a subject matter problem.

In fact, it is difficult to overestimate the importance of the homogeneity problem. Statisticians of many different schools are aware of it. For example, Chuprov (1959) in 1910 wrote, "the move from relative frequencies to mathematical probabilities can be carried out only on the basis of a homogeneous data set". Boldrini (1972, p. 162) also underlines the importance of homogeneity in statistics. He wrote, "... an experimental probability is an induction; it is based on homogeneous factual data... ".

Referring to the quality of statistical information, Fisher (see Fisher, 1959) wrote, "The statistician is no longer an alchemist expected to produce gold from any worthless material offered him". The potential consequence of employing non-homogeneous data is the complete invalidation of statistical inferences based on them. Therefore, this author entirely agrees with Pearson and Hartley (1954, p. 83) that "it is the function of statistical methods to emphasise that precise conclusions cannot be drawn from inadequate data".

Unfortunately, statisticians do not have statistical methods for testing the data set homogeneity, and consequently they are in the position of having to produce gold from statistical data of unknown quality. The bottom line must be: our statistical inferences are only as good as the data on which they are based.

The solution of the homogeneity problem lies in probability theory. The theory of probability is the foundation of statistics. It traces effect from cause by employing the homogeneity principle. Regarding this principle, Khinchin (1961) wrote: "The homogeneity principle in probability theory is a fundamental and unique doctrine employed for theoretical predictions of probabilistic events in specific cases".

The homogeneity principle in probability theory is associated with the notion of the homogeneous population. The population is considered to be homogeneous if its events have equal chances of occurring. Hence, the main characteristics of the homogeneous population is the constant success probability of its elements.

Probability theory employs the deductive method of thinking, going from cause to effect, i.e. from a homogeneous population to samples. However, these samples may be representative or non-representative of the homogeneous population. The move from effect to cause is possible only if the sample is representative of the homogeneous population. This move is built upon inductive reasoning, formulated in terms of probability theory. The first requirement of inductive reasoning is that the population is homogeneous. The second requirement is that the sample is representative.

Inductive reasoning is the essence of the New Quality Philosophy in statistics (NQP). The NQP studies inward manifestations of homogeneous mass phenomena. It considers statistics to be a science and defines it as follows: statistics is a universal methodological science which quantifies the cause-effect relationship. This quantification is valid only when the sample is representative of the homogeneous population. According to this definition, statistical research must consists of two stages: firstly the evaluation of the sample quality, and secondly, the construction of the statistical model.

4. Axioms of the New Quality Philosophy of statistics (NQP)

The New Quality Philosophy of statistics (NQP) is built upon four axioms which are formulated in terms of probability theory.

The First Axiom

The homogeneous population is known *a priori* as a homogeneous invisible population (HIP). It consists of two groups of homogeneous invisible causes, successful (S) and unsuccessful (U). The successful causes exert a positive influence on the event. The unsuccessful causes exert a negative influence on the event. These two groups of invisible causes are homogeneous if two conditions are met. Firstly, each cause must have an equal chance of occurring. Secondly, each group of causes has its own constant influence on the effect.

Thus, the first axiom is that the HIP is a Bernoulli probability distribution in terms of homogeneous invisible causes: successful and unsuccessful.

The Second Axiom

The success probability (P) of the HIP is the proportion of the number of successful causes in the total number of all causes. It is considered to be an explanatory factor. Under the influence of this factor, the average effect is generated. The visible average effect is a mean of the statistical population which has to be representative of the HIP.

Thus, the second axiom is as follows: the mean of the HIP is equal to the mean of the representative statistical population.

The Third Axiom

The average effect, generated under the influence of the homogeneous invisible causes (HIC), is presented by the mean of the representative statistical population. The individual effect, generated under the influence of the HIC, is presented by a statistical datum. The statistical datum is less than the mean under the influence of unsuccessful causes. It is greater than the mean under the influence of successful causes, and it is equal to the mean under the influence of both causes, unsuccessful and successful. Thus, combinations between the HICs determine the size of a statistical datum with respect to the mean of the statistical population when it is representative of the HIP. These combinations are the following:

$$\begin{array}{lllll} U \cap U & = & x_1 & < & \bar{X} \\ S \cap S & = & x_2 & > & \bar{X} \\ U \cap U & = & x_3 & = & \bar{X} \\ S \cap U & = & x_4 & = & \bar{X} \end{array}$$

Therefore, a statistical datum can be considered as a combined event or as an indivisible sample with the corresponding probability. Probabilities of combined events can be computed as soon as the success probability of the HIP is known. This means that each datum has an individual probability or, according to Markov's terminology, "a single case" is associated with its probability (Markov, 1911).

Thus, the third axiom is that a statistical datum is a combined event, or an indivisible sample, originated under the influence of the HIC.

The Fourth Axiom

The HIP is characterised by the variable, consisting of two types of homogeneous invisible causes, and by the success probability (P). The set of homogeneous invisible samples (HIS), drawn from the HIP, is characterised by the variable, consisting of combinations between the HIC, and by probabilities of this variable.

The size of combinations between the HIC must be equal to two in order to meet requirements of the principle of minimum and the principle of representativeness. The requirement of the principle of minimum is met when the size of the HIS is equal to the number of causes of the HIP, that is, two. The requirement of the principle of representativeness with respect to the HIP is met when the number of representative combinations between the HIC is equal to 50% or more.

For example, we expect the following combinations between the HICs when homogeneous invisible samples are drawn from the HIP:

$$U \cap U, S \cap S, U \cap S, S \cap U$$

These combinations are combined events (CE) or the HIS. The set of the CE consists of two non-representative combined events (NCE)

$$U \cap U, S \cap S,$$

and two representative combined events (RCE)

$$U \cap S, S \cap U.$$

Let us designate the NCE by 0 and the RCE by 1. Then our set of the CE in terms of the theoretical values of representativeness (TVR) can be written as follows:

$$0, 0, 1, 1.$$

This set of TVR (STVR) is considered to be a representative set of the CE (RSCE) with respect to the HIP because it contains 50% of the RCE.

Our STVR is a variable of the representative set of combined events. Probabilities of this variable are computed with the help of the multiplication formula for independent events (causes) when the success probability of the HIP is known. This probability distribution of the TVP is representative of the HIP because two conditions are met:

1. The variable of the CE is a representative set of the HIS.

2. The success probability of the CE is constant. It is used for computing probabilities of the variable of the CE. This means that the probability distribution of the CE is generate under the influence of the explanatory factor, which is a success probability of the HIP.

Since each statistical datum is a combined event, the fourth axiom can be formulated as follows: the data set is good if it is a representative set of the combined events generated by the HIC.

Bibliography

Boldrini M. (1972), *Scientific Truth and Statistical Method*, Hafner Publishing Co. New York.

Chuprov A.A. (1959), *Essays on the Statistics Theory*, Gostatizdat, Moscow (in Russian).

Fisher R.A. (1959), *Statistical Methods and Scientific Inference*, Oliver and Boyd Ltd., Edinburgh.

Khinchin A. (1961), *The Frequency Theory of R. von Mises and Modern Ideas of Probability Theory*, Problem of Philosophy, #1, pp. 91-102; #2, pp. 77-89 (in Russian).

Markov A.A. (1911), On the Basic Principles of the Calculus of Probability and on the Law of Large Numbers, in *The Correspondence Between A.A. Markov and A.A. Chuprov on the Theory of Probability*, Springer-Verlag, New York (1981).

Pearson E.S. and Hartley H.S. (1954), *Biometrika Tables*, Vol. 1, 3rd ed., The University Press, Cambridge.

Sachs L. (1982), *Applied Statistics. A Handbook of Techniques*, Springer-Verlag, New York.

Shvyrkov V.V. (1992), *New Thinking about Quality Control*, Throwkoff Press, Santa Rosa, CA.

Shvyrkov V.V. (1993), *Axiomatic Principles of Statistical Science*, Throwkoff Press, Santa Rosa, CA .

LEARNING METHODS INVOLVING COMMUNICATION BY STUDENTS

John Taffe
Swinburne University of Technology
P.O. Box 218, Hawthorn 3122, Australia

1. Introduction

Students learning statistics are really much happier when there is a story, or storyline which forms the background for their enquiries.

I wish to maintain also that they derive great benefit (and not only to their study of statistics) from explaining to their peers and teachers what they have done to solve a problem.

The component of our 1st semester, 1st year statistics course in Applied Science that I discuss here incorporates both of these features, but I have chosen to concentrate on the latter.

I have proposed elsewhere (Taffe, 1990) that the main beneficiary of any explanation is the person who gives the explanation. The process of reorganising ideas into a linear chain of description and explication, in real time, helps you to see new connections and new perspectives on the content as a whole. You increase your mastery of the content in leaps and bounds.

For the listener, however, explanations often mystify. Recall Bock's Law: If it can be misunderstood, it will be, and if it can't be misunderstood, it will be; and Taffe's Corollary: Hardly anyone will understand any explanation, no matter how clear.

We may conclude from these laws that the well known and highly esteemed 'gravity feed' method of teaching may not be optimal in terms of student learning, at least not if used at all times.

At Swinburne, first year Applied Science students have the opportunity to transfer the benefits of explanation to themselves.

2. Communication by students

The Presentation. For the first semester unit "Data Analysis and Probability" there are two assignments, which are really aspects of the same single assignment: the Presentation and the Written Report. They are based on a small set of investigation-type problems, which provide the storyline (see appendix).

Students are given to understand that their task is not to solve the problem - they will be given any help they require if the problem proves difficult. Their task is rather to communicate to their teacher and some of their peers the results of their work on the problem, and to answer questions about it.

Two weeks after being given the problem they present their communication in 15 minutes (including 5 minutes for questions). This is assessed by their teacher for evidence that they understand (even a small part of the original problem), and for their success in communication.

One week later they hand in a written report based on their talk and preferably answering some of the questions raised in discussion at the time of the talk.

Guidelines. The students are given brief guidelines (see appendix) as to what is expected in both parts of the assignment.

Discussion. It is well known that what 'drives' students (as far as their academic work is concerned) is the assessment of the work - what is assessed and how it is assessed. By including in the course a component (30% combined weighting for the presentation and the written report) of work which is assessed for quality of communication we indicate that we think this an important part of their statistical education.

Generally the students enjoy the work, and the problems with which they become familiar can provide a basis for exemplifying later theory. I recommend assignments of this kind to you.

Bibliography

Taffe J. (1990), The Gravity Feed Model of Teaching, Laws and Strategies, in D. Vere Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics,* International Statistical Institute, Voorburg, Vol. 2, pp. 43-50.

APPENDIX

Swinburne University of Technology
Department of Mathematics
SM184 Data Analysis and Probability 1993

Guidelines for presentations

Presenter

Your presentation should be along the following lines:

Introduce yourself formally.

State the central idea or question you are investigating in an easily understandable way. You should say how you expected the investigation would turn out before you started it.

Give an outline of your talk in the form of a few headings, to let your audience know what to expect.

Concentrate on illustrating 2 or 3 main ideas – don't pay too much attention to details.

State your conclusions and invite questions.

Things to keep in mind

Time – be strict with yourself. Allow 10 minutes for your talk and 5 minutes for questions and comments.

Illustrations – on OHP slides, on the blackboard. They should be simple – one main illustration on each. Writing should be large and clear.

Speak more slowly and louder than you usually do. Look around at everyone in the audience (not just at your lecturer).

Audience

Being in the audience is just as important as giving the presentation. Take seriously your responsibility to ask questions and to offer comments. Try to imagine working on this topic yourself and think how you would have handled it. This will help you to think of questions and comments.

Assessment

In assessing your presentation marks will be given for:

Thorough preparation
Clear sequence of ideas
Appropriate techniques and illustrations
Drawing appropriate conclusions
Handling of questions

A mark out of 15 will be given.

Guidelines for written reports

The written report should be constructed on the following lines:

Introduction

Say what the topic asks you to do, discuss the central idea and how it relates to the study of statistics. Say what kinds of results you expected before you started to investigate the question. 3 marks

Method

Set out clearly how you went about the task, what decisions you had to make, what difficulties you met and how you resolved them. 3 marks

Results

Present your main results in graphic or tabular form, and point out the main things which should be noticed. 3 marks

Conclusion

Set out what conclusions you have come to in the light of your results. If possible comment on how the topic relates to parts of your statistics course and to real-world problems. You may wish to make suggestions about improving the topic for future students. 4 marks

Appendix

You should include a listing of the data at the end.

The whole report should take no more than 3 o 4 pages. Two more marks will be awarded for overall presentation and quality of the report. These will be awarded as follows, according to your lecturer's judgment:

0 The report fulfils the basic requirements.
1 The report shows some evidence of original thought, or perhaps extends the topic in some way or makes a useful comment.

2 The report does more than one of the above-mentioned things and is excellently presented.

The report should be handed in one week after the presentation.

TOPIC 1 PSSST!

In a group of nine friends, one starts a rumour. She tells it to one of her friends, and the friend passes it on to the next member of the group that he or she meets. This continues until the rumour is told to someone in the group who has already heard it. This person does not pass it on, so the rumour stops spreading. What is the probability that everyone in the group hears the rumour? Use random numbers from a table or from your calculator or generated by Minitab to simulate this process and give an estimate of the probability.

TOPIC 2 BIRTHDAY SNAP!

Ask students coming out of the engineering building (or going in to the library or the café ...) what day of the month their birthday falls on (possible answers are 1, 2, 3, ..., 31). Write down the answers in the order they are given. When a number is repeated, say 'Snap!', rule off the list and start again. Do this until you have a large number of snaps (at least 25). Count how many people were required for each snap, and summarise the sample distribution of this number in a suitable way.
Now consider this problem: How big does a group of people need to be for at least an even chance (probability ≥ 0.5) of a 'birthday snap'? How can you use your sample to estimate this number? Can you suggest other ways of solving this problem?

TOPIC 3 BUFFON'S TOOTHPICK

Obtain a large sheet of paper (e.g. butcher's paper) and a toothpick. Measure the length of the toothpick and rule up the paper with parallel lines whose distance apart is the same as the length of the toothpick. (Alternatively, cut a stick so that its length is the same as the width of the boards in a wooden floor). Throw the toothpick onto the paper (or the stick onto the floor) and note whether it lands across one of the lines. After the fifth throw note down the relative frequency of 'crossings'. Do this again after the tenth, fifteenth, twentieth, and so on till the hundredth throw. Plot these relative frequencies against the number of throws. Discuss the results. The probability of a crossing has been caculated as $2/\pi$. Is this consistent with your results? (This is usually referred to as Buffon's Needle Problem).

TOPIC 4 A DICE PROBLEM

Is it easier to throw (at least) one six with six dice or (at least) two with twelve? Borrow some dice from your lecturer and investigate this problem. Use 100 throws with six and 100 with twelve. In each case record the relative frequency of successful throws after the fifth, tenth, ..., hundredth throws and plot these relative frequencies against the number of throws. Discuss the results and predict what would be likely to happen if three or more sixes had to be thrown with eighteen dice.
How could Minitab be used to simulate the rolling of the dice?
Is there a way of calculating exact answers?

TOPIC 5 A SIMPLE CODE

Select a page from a novel and use it to construct a frequency distribution of the 26 letters of the alphabet. Compare this frequency distribution with the distribution of letters in the coded message below. Explain how comparing letter frequencies in this way can be useful in deciphering codes, and translate the message:

Wbw rhn zh mh max Vhfxwr Yxlmbote? B ptl mhllbgz ni uxmpxxg 'Max Mpbml" tgw 'Phz-t-ktft', unm ymgteer pxgm mh 'Max Zbke'l Zhmmt Xtm' tgw etnzaxw frlxey lbvd!

TOPIC 6 POKER: THE CHANCE OF A PAIR

Deal 5 cards from a well shuffled pack of 52 cards. Record whether this hand contains one 'pair' (and nothing else) or not. Repeat this experiment 20 times and estimate the probability of being dealt a hand containing one pair.
Using random numbers (and possibly Minitab) simulate 100 poker hands and use the results to estimate the probability of one pair.

Calculate the probability and comment on the accuracy of the two simulation methods.

Discuss the problem of estimating the probability of 'three of a kind'.

TOPIC 7 HIEU TAKES A PUNT

Hopeful Hieu is soon to do a multiple choice test. The test is to consist of 20 questions, each with 4 possible answers. Only one of these four answers is right.

Hieu is worried. His social life has cut into his study time quite a lot. If he knows the answers to 9 of the questions he is confident of passing, because he would need only one of 11 guesses at the other questions to be correct. But what if he knows only 8, or 7, or 6, or...? How many does he need to know to have at least a 50% chance of passing? Could his chance of passing drop below 5%?

Statistics is not his best subject. Help Hieu by using random numbers to simulate his guesses (a random number less than 25 means a correct guess). Suppose he knows 9 answers. Simulate an attempt at the test by making him guess until he gets another answer right or all 11 wrong. Do this 20 times and record his percentage of 'passed'. (This is your estimate of his chance of passing when he knows only 9 answers). Then suppose he knows 8, and so on.

TOPIC 8 HAT'S ALL FOLKS!

One evening at the theatre ten hats are handed in at the cloakroom. The attendant is so busy she forgets to put tickets on the hats. When the hats are called for after the performance she simply hands them out at random. The patrons are so engrossed in their discussion of the performance that they accept the hats without checking that they have the right ones. How many hats can we expect to be returned to their rightful owners?

Number the ten hats owners 0 to 9 and use a table of random numbers or a calculator to assign the hats randomly. Repeat this procedure until think you have a reliable estimate of the expected number of correct returns. After each simulation calculate the average number of correct returns per simulation and graph this against the number of simulations.

TOPIC 9 SITTING DUCK

Each day a hunter goes duck shooting. She fires repeatedly until she shoots a duck and then she takes the duck home to her hungry family. The shots she fires are independent and the probability of hitting a duck with each shot is .2. Her husband, concerned at the expense of his wife's wasted bullets, finds out that free range ducks can be bought from the local market for $4.50 each and suggests that since bullets cost $1 each it may be cheaper to buy the ducks. Use Minitab to simulate 100 days' shooting and decide whether the husband is correct. How could you have answered this question without using Minitab?

TOPIC 10 BLOOD GROUPS

Suppose that in a large population the blood types A, B, AB and O occur in the proportions 30%, 20%, 10% and 40% respectively. From this population, four blood donors are selected at random. By simulating 100 such samples of four people, estimate the proportion of the population that would, on average, not have their blood type supplied from the 4-persons sample. Compare this estimate with the exact value of the average proportion uncatered for by the four-person donor group.

TOPIC 11 TESTING, TESTING!!

Blood samples frome the 11 members of a cricket team are taken, and are mixed together. If this mixture contains any trace of the performance enhancing drug Spinoutaside, each of the 11 players will then be tested separately. This means that either one test or 12 tests will be required. Assume that the probability of each player using Spinoutaside is 0.1, and that the players' drug-taking (or non drug-taking) habits are independent of each other. Estimate the expected number of tests required, by carrying out 100 simulations. Compare this estimate with the exact expected number of tests required.

TOPIC 12 GOODY TWO SHOES

The lights have gone out. You have 10 pairs of shoes in the bottom of your wardrobe, but they are in an untidy pile. You pull out six shoes. What is the probability that you have at least one pair? Use a simulation method to estimate this probability.

TOPIC 13 FLYING DUCKS

It is early morning on the first day of the duck season. A flight of 10 ducks takes off. Ten hunters shoot and each hits the duck he aims at. How many ducks escape?
Use random numbers to simulate this situation 10 times to get an estimate of the expected number of escaped ducks. Do this again 50 times and then 100 times and comment on how the reliability of your estimate improves. Estimate the probability that no more than 3 ducks escape.

TOPIC 14 TWO PEAKS

The town of Two Peaks has an unusual legal system. The procedure in any trial is as follows:

(i) The judge rolls a die and reports the number, k, to the 4 jurors.

(ii) Each juror uses his or her calculator to generate a random number between 0 and 1, and votes to convict the accused if, and only if, the first digit of the number is ≤ k.

(iii) The accused is found guilty only if all four jurors vote to convict.

Leland De Palma is on trial in Two Peaks for the murder of his mother-in-law Lorna.
Using a die and your calculator, estimate the chance of him being convicted by simulating 100 court verdicts. Calculate the probability exactly and compare the answer with your estimate.

THE IASE AND PROBLEMS OF STATISTICAL EDUCATION IN DEVELOPING COUNTRIES

Invited paper David Vere-Jones

THE IASE AND PROBLEMS OF STATISTICAL EDUCATION IN DEVELOPING COUNTRIES

David Vere-Jones
Institute of Statistics and Operations Research
Victoria University, P.O. Box 600, Wellington, New Zealand

1. Introduction

I have chosen this theme, despite the near impossibility of treating it in anything beyond platitudes, for two main reasons.

The first is personal. When I first came into the role of Chairman of the ISI Education Committee, it was after several years of effort to build up a self-sustaining East Asian and Pacific Regional Committee of the Bernoulli Society, essentially a regional statistical grouping. This brought me into direct contact with statisticians in many Asian countries who were trying to establish the role of statistics in their educational programmes. I had hoped to promote similar ends through the Education Committee, but addressing the problems of statistical education in the developing countries proved to be as intractable from this vantage point as it had been earlier. Still, I do not want to leave my present position without trying to draw something from these efforts.

The second reason is a more practical one. In the past the ISI Education Programme had as one of its earliest and continuing motives the need to increase the number of statistically trained staff for governments in the developing countries. Out of its responsibilities in this direction derive both its UNESCO grant and its special link with the International Statistical Education Centre in Calcutta (ISEC). A crucial policy question for the IASE is how far it will wish to pick up responsibilities of this kind from the old Education Committee. The Association has come into being, as I see it, primarily to represent the interests of a newly recognised professional group, the teachers of statistics. Through its historical links with the ISI, and its continuing membership of the ISI family, there comes an additional responsibility, to further the educational objectives of the parent organisation. The two directions are not necessarily opposed, but as a member of the ISI family the IASE will face issues that would not have come its way had it been set up as an independent professional body outside that family. Insofar as both the UNESCO funding and the ISEC link are currently in doubt for other reasons, the questions are urgent as well as fundamental. Indeed, they are already down for discussion at the Florence Meetings, and the attitude taken by the IASE is likely to be crucial to the way in which they are approached by the parent body.

In this paper I shall try to join together elements from both motivations. The next section (§2) is a summary of a more extended history of the ISI Education Committee that I have given recently (Vere-Jones, 1993). Its purpose here is to show how the ISI Education Programme itself developed, and something of its immediate objectives at the time the IASE was formed and the old Education Committee disbanded. In the following three sections (§3-5) I have tried to represent, in a very generalised way, some of the characteristic difficulties of "furthering" statistical education in the developing countries. While doing so I have noted some possible strategies that the IASE might take in trying to overcome them, and in the final section I have collected these suggestions together.

2. The historical legacy

The ISI Education Programme has been reviewed, critically and otherwise, on at least five previous occasions (Nixon, 1960; Goudswaard, 1964; Zarkovich, 1976; Gani, 1979; Gani, 1987), listed in detail from bibliography, and also in Vere-Jones (1993), which attempts to bring the record up to the date when the IASE was formed. I summarise a few key points here to provide a basis for considering the role of the IASE in relation to the ISI's own educational objectives.

Up until the time of the Second World War, the ISI had something of the character of a club for government statisticians, "a semi-official organisation, collecting international statistics for government use, drafting conventions and the like" (Nixon, 1960). Its membership consisted of the government statisticians, ex officio, and a small number of highly distinguished scholars and civil servants from a broad spectrum of occupations loosely related to statistics: economics, agriculture, the social sciences, mathematical statistics, etc. From its inauguration in 1885 education had figured among its concerns, but not in a systematic way.

In 1948, with the formation of the United Nations and its agencies, Dr Stuart Rice, then President of the ISI, called for a reexamination of its objectives and a reorganisation of its functions. He argued forcefully that statistical education should be a major function. "Relieved of its direct responsibilities for inter-governmental action, the Institute is now freed for ancillary services of an advisory or contractual character to the United

Nations, its specialist agencies, or member governments. The proposed international programme for education in statistics is of this character". (Rice, 1948). Changes to the ISI Statute which, among other things, recognised this increased role of statistical education were implemented in 1948, and a major paper by Rice, on "The Furtherance of Statistical Education" (Rice, 1949) was adopted by UNESCO and the ISI as the basis for future development. Among the specific objectives proposed at that time were the creation of international statistical training centres, development of a fellowship programme, dissemination of reference materials, a survey of existing facilities, etc.

A key reason for the support of UNESCO and the UN agencies was the recognised need to increase the number of statistical staff in developing countries. Such staff were required to furnish appropriate statistical summaries for UN purposes. As an operating responsibility, statistical education was regarded as a function of the UN; the ISI was recognised as the preferred agency for promoting statistical education throughout the world.

The Education Committee itself was also set up in 1948, first meeting in 1949. Rice himself was the first chairman, a role which he continued to carry even after his term as President of ISI had ended. He was succeeded by the famous Indian statistician, Mahalanobis, who was also Director of the Indian Statistical Institute. Through his influence, the first ISEC was set up in Calcutta in 1950, with funding provided by the Indian Government and by UNESCO through the ISI. Later the funding was largely taken over by the Indian Government, one part of the funding being channelled through the ISI for the selection and support of visiting teachers. A second centre was opened in Beirut in 1953, and several other centres have opened since then, but the ISI has not had direct links to these centres.

The funding from UNESCO, initiated in 1949, has continued to the present time in the form of an annual subvention for the promotion of statistical education. The Indian Government funding for ISEC Calcutta via the ISI has also continued. However both grants are now under serious threat. The Statistical Division of UNESCO is no longer able to continue the subvention at its earlier rate of around $ 20,000 per annum and may be obliged to cut it completely. Payments from the Indian Government have been extremely irregular and the ISI Executive has set up a subcommittee, to reexamine the basis for these arrangements. In both cases the IASE is likely to be plunged into negotiations which may have a crucial influence on both is own and the ISI finances, and on the direction of its future activities.

Returning to the Education Programme, with the passing of the years, its emphasis shifted from training statistical staff in developing countries towards the promotion of statistical education in universities and schools. These changes reflected both changes in the ISI membership, which has now come to be dominated by academic and professional statisticians, and the growing prominence of statistics in university and school programmes. They were particularly pronounced in the final phase of the Education Committee, which in Vere-Jones (1993) I date from Zarkovich's reappraisal (1976) of the Programme, and covers the many initiatives taken during Joe Gani's energetic chairmanship of the Committee (1979-1987).

Among the most important aspects of its recent activities, I would place first the continuing programme of international meetings, both special-purpose Round-Table Conferences, and the series of ICOTS Meetings, of which the 4th is scheduled to be held in Marocco in 1994. This programme was directed by the Committee's "Task Force on Conferences", led initially by Lennart Råde and more recently by Kerstin Vännman, and has been carried on by the Interim Executive pending its further evolution under the auspices of IASE.

Another important aspect has been in publications. One of Gani's initiatives was the implementation of the two small books on "Teaching Statistics Around the World", one in schools and the other in universities and technical institutes (Barnett, 1982; Loynes, 1987). The "Totsas" newsletter entered new territory in providing for the first time a newsletter specifically directed at teachers of statistics at school level. Currently, newsletter material from IASE appears in two places, the centrefold in the journal "Teaching Statistics", and the IASE section of the overall ISI newsletter.

I see no real difficulty in transferring these and other initiatives taken by the Education Committee in recent years to the IASE and its Executive.

The one aspect which does not seem to me to have been fully thought through is the fundamental one relating to the educational objectives of the ISI itself. It is clear that setting up an international association for statistical education is entirely consistent with these objectives. The difficulty comes in the need to clarify to what extent and by what mechanism the ISI intends to delegate to the IASE its overall responsibilities for statistical education. As we have seen, these form a major part of the ISI's own objectives. The IASE is in the first instance a professional organisation representing the interests of those concerned with Statistical Education. It is not immediately clear that such an organisation is the most appropriate body for discharging the ISI's own responsibilities. Such a debate would hardly have arisen if the IASE had been formed outside the ISI family. Being inside the family, the IASE inherits certain family obligations, which may well make its role richer and more interesting, but also more difficult.

At a practical level, the IASE Executive may need to initiate discussions with the ISI Executive, and also with the other international organisations in the ISI family (IASC, IAOS, IASS and the Bernoulli Society) as to how is should seek to take up, and then discharge, these more general responsibilities. Since these other sections may well have educational programmes of their own, the IASE could consider starting a cycle, aimed at holding one joint meeting each year with one of the other sections. Projects of an educational character could also benefit from joint sponsorship by IASE and one of the other ISI sections; for example, projects relating to the training of statistical staff could well be taken up jointly with IAOS or IASS.

In the past, representatives of the UN agencies most closely concerned with the ISI Education programme had ex officio status on the Education Committee. The IASE Executive has suggested that they be invited to accept observer status on the IASE Executive, with the right to receive agenda notices and minutes, and to contribute to discussions of policy issues.

3. The Developing Countries: some general cautions

Although my main purpose is to consider ways the Association may assist in the promotion of statistical education in developing countries, it may be appropriate to start off by cautioning against too simplistic an approach. It is important to appreciate some of the major reasons why progress has not been easy in the past, and may be just as hard in the future.

The first important point is the extraordinary diversity among the developing countries, and their individual complexities. Doyle and Brown (1991), in their vivid account of the problems facing statistical trainers working with the South Pacific Commission, assert bluntly that "the biggest difficulty encountered in providing statistical training to the SPC's member countries can be summed up in one word – diversity". The Pacific Islands may evoke a uniform image of tranquil atolls, but the reverse is the more common rule. They are diverse, and individually complex. For some years I had some contact with Fiji as an external assessor in mathematics for the University of the South Pacific, visiting in that capacity not only the main campus in Suva but also its offshoots (now independent institutions) in Western Samoa and Tonga. Talking with staff about the students' performance in different subjects, one became dimly aware of the difficulties that they faced, the totally differing stocks of experience, expectation, mental images, reasoning patterns, etc., that they brought with them to their studies; of the cultural minefields that could be set off by the most innocent of statistical examples; of their prospects for employment in an environment dominated by political intrigues, social and racial pedigrees, large discrepancies in wealth and influence. And I am sceptical that Fiji is exceptional in this regard; rather, I suspect, representative of what one might find in many smaller countries of Africa, South America, Asia, the Middle East.

In such a context, one should not expect to find the arguments in favour of increasing the statistical component in the university or, more particularly, the school programme, as persuasive as they seem to us in the West. Even in Western countries, this is a recent development, none too secure in many cases, the current end-point of a long evolutionary process in mathematical education. In Vere-Jones (1993) I have argued that it represents a break from an elitist view of the role of mathematical education to a more democratic one. The developing countries are only just starting this evolutionary path, and their routes may end up taking quite different directions, or they may need to wait as long as we have for the arguments in favour of change to outweigh more immediate concerns of economic development or personal advancement within the existing system.

Nor can one afford to overlook the political climate. Florence Nightingale was one of the first to recognise the force of political arguments backed up by hard statistical facts. Unfortunately this is no guarantee that the statistician has an influential political role. More likely it will be just the reverse. The statistical profession is likely to be at the bottom of the heap, well below the economists and maybe even the scientists, though influential jobs for scientists are not a common phenomenon among developing countries either. Nor do all regimes want to be reminded of hard statistical facts, however salutary they may seem, and necessary for future national progress.

Two of the ICOTS 3 papers present eloquent pleas to recognise the political impact of statistics, though from very different points of view. Tulya Muhika (1991), looking towards a renaissance of statistical services in Africa, decries the influence of Western style economists on African governments, in frequently leading them away from policies of benefit to their own people. He sees the quality and integrity of a country's statistical services as essential to its progress and long-term viability.

Jain (1991), in her plenary address emphasises the danger of bias in the selection of indices used to describe a country's economic and social situation. Current usage of Western countries, and also international organisations such as the World Bank and the IMF, emphasises cash transactions and commercial indicators; in doing so they may force the pattern of development into the hands of Western business interests and away from the interests of the local people. In particular the value of womens' work, frequently unpaid, may be grossly undervalued in this process.

Just how this impacts on statistical education is not clear, but it may be as well to admit that statistics as a school and university subject is not so politically neutral as one might assume.

Finally one should bear in mind the limitations of the IASE itself. Its strengths, like those of the ISI, are likely to derive from its ability to bring together leading educators from many different regions, to examine problems, suggest remedies, initiate new developments – what Joe Gani (1987) jocularly and almost disparagingly refers to as the "gadfly" role. Yet it is just this role, in the seeding of new developments, and the timely identification of potential problems and new issues, that it may have greatest value. When it comes to the operational role, on the other hand, its possibilities are severely limited by its being largely dependent on the voluntary efforts of a dispersed membership with other preoccupations. Its lack of a physical base (the ISI Permanent Office in the Voorburg is a modest suite of rooms in the building of the Dutch Statistics Bureau, which it occupies by courtesy of the Bureau) and lack of significant financial resources are further problems. For funding, IASE is almost totally dependent on grants and the minuscule revenue generated by members' dues. Even as a lobbying group its resources remain to be developed.

Notwithstanding these difficulties, there are bound to be opportunities to set useful actions in train if the Association is poised to take advantage of them. Like the Boy Scouts, it has to "be prepared" for opportunities turning up. In this regard it might consider setting up a group, like the Education Committee's Task Force on Conferences, to receive or canvass suggestions for assisting the developing countries, to assess the suggestions, and to pass them on to the Executive with considered recommendations.

4. Working with the least developed countries

As mental preparation, picture a small, probably tropical, country, where the material standard of living is very low for the great majority of the population; where education, even at primary level, is not universal, while passing into a secondary school represents a substantial social and financial step; where a small elite captures most of the wealth, and controls much of the government apparatus; where local money has no significant exchange value, so international travel is impossible for most of the population; where communications are slow, fax and E-mail facilities exist only in one or two centres, and are in any case unreliable.

Where does statistical education impinge on such an environment, and in what ways could the IASE, or any similar organisation, become involved?

I shall mention three possibilities.

In the first place, it is just in such small countries that problems of recruiting and retraining suitable statistical staff are at their most acute. It is likely that local "country courses", such as those described by Doyle and Brian (1991), are run there from time to time by the nearest International Training Centre, or by other inter-governmental organisations with their own interests in seeing that at least the most basic statistical data is compiled.

Putting on such courses is a highly professional task, and IASE's first step could be to investigate whether IASE membership could be of value to the staff involved in such work. It might also consider approaching the International Training Centres themselves to see whether they might join as Institutional members. In the past, the ISI education Committee has from time to time attempted to bring together the directors of International Training Centres for joint meetings and exchange of views. It has also published lists of institutions providing courses suitable for trainee staff from developing countries, though my impression is that little use was ever made of these. However, since such training activities are managed through UN or similar agencies, IASE/ISI does not have any clear role in their work. It starts off on the outside and will need to show itself to be of value and relevance. My impression is that there is some need for international links at the professional level, so that a vehicle to foster such links, giving staff the opportunity to meet and exchange experiences and evaluate different approaches, could be of value.

A second possibility for the IASE concerns the development of basic teaching material, at school level or higher. During the ICOTS 3 Meeting, Nasim Ahmad, at that time the headmaster of a mssion schools in the Gambia, outlined a proposal for developing a common basic text for high schools in a group of neighbouring anglophone countries. As he explained, the cost of a regular Western text might be the equivalent of several months' wages to the parents of a child going on to secondary education. The need was for cheaper material which gave more prominence to statistical concepts within the secondary mathematics curriculum. Statistics recommended itself not only for the usual reasons of practical relevance etc., but also because junior statistical staff were commonly recruited directly from school and so were helped by some contact with notions of economic indices, survey methods, etc. Although the idea clearly had merit, it was not easy to see whether IASE/ISI could do anything to help.

It seemed to me at the time that three basic components had to be linked together. The first was a group of local teachers or educators backing the idea and with local knowledge of what was needed. The second was a physical location where one or more from the group could have access to advice and library resources, and could spearhead the writing task, while taking time off from their regular teaching duties. The third was access to a funding body willing to cover the costs of these preparations. These are, of course, familiar problems to anyone developing statistical education elsewhere in the world, but made more difficult by the lack of local funding and expertise, and the special problems of producing low-cost material, either in English as a second language, or in the vernacular.

The historical record suggests that such tasks are not easy for a body such as IASE to undertake. The fate of the ISI Video Project (see Bryson, 1991) provides a salutary lesson in this regard. This was a strongly-backed scheme for using a video approach to reduce the costs of training primary staff (interviewers etc.) for a nutritional survey. It successfully passed a pilot study stage, but failed at the substantive stage through the inability to find a relatively small residual amount of funding. Perhaps the IASE should take one step back from direct involvement in such projects and restrict its efforts to providing support at a more personal level to the teachers and others already engaged in such efforts. However, I would be reluctant to rule out the possibility of a more direct involvement.

A third possible route for the IASE is through support to individual staff in universities or other teaching institutions, who can be encouraged in their attempts to initiate teaching and curriculum change. Again with ICOTS 3, we obtained support from the Commonwealth Foundation to bring a 2-person team of "key teachers" from selected commonwealth countries to New Zealand, to participate in the conference and take part in an auxiliary programme of lectures and visits to New Zealand schools. Unfortunately the funding came very late, and we met insuperable difficulties in trying to locate suitable teachers at a distance. We did not want to be fobbed off with government-selected educational bureaucrats, and ultimately had to give up the attempt.

While such examples are not encouraging, the more IASE can do to recruit members from the developing countries, the better chance it will have of working out more effective strategies. In this regard the range of countries already represented on the IASE membership list is highly gratifying, and an important task is to keep their needs very firmly in mind. In the long run it has to be through individuals, on site, working within the local context, that real progress will depend.

5. Working with countries in transition

Here I have in mind a wide spectrum of countries, from many different parts of the world, which are evolving rapidly, in some cases even surpassing Western countries in terms of economic growth. A typical country in this category is likely to have substantial historical and cultural traditions of its own; a number of universities, including one or two of international standard; a widespread programme of mass education, reaching even into the rural areas at primary level, and with a broad entry into secondary education. Even so there is likely to be great disparity of wealth between rich and poor, with racial and economic tensions leading to unstable political regimes.

My experience of working with the Asian countries suggests that while resources remain an underlying background problem (but where is this not so?), the most difficult obstacles may be cultural and attitudinal ones. Here also, there is little that is not present in the Western countries, but the problems are more acute because of economic pressures and political instability.

To give a rather hackneyed and probably unfair example of possible cultural problems, Asian cultures are often accused of a rather "mandarin" attitude to practical work. Social standing is demonstrated by how far one can remove oneself from the need to undertake physical labour. Since statistics is at root a "hands-on" practical subject, in trying to promote statistics one may find oneself struggling against an uphill cultural gradient. Pure mathematics and even probability theory may have a more attractive image, while statistics remains a poor man's option.

Whether or not this is really a factor, in Asian countries, as elsewhere, it is very common to find a traditional, non-statistical mathematics programme firmly entrenched within the upper school, and especially in the university entrance examinations.

It is at this point that economic factors start to play a role. In many countries in transition, entry into a university programme is the critical step out of a poverty trap in which families may have suffered for generations. Entry is therefore fiercely competitive, far more so than in most Western countries, where progression beyond the secondary level is becoming the norm rather than the exception. Statistics in the schools, at least as it is now conceived in the Western countries, is not intended as a vehicle for selecting the elite. It forms a major plank in a new programme of "mathematics for all". In a highly competitive examination environment one is faced with a dilemma. Either statistics is taught as an academic subject, suitable for filtering out the selected few from the masses, or it drops out of the syllabus for the entrance examination and thereby loses its ability to influence the school programme.

In such an environment progress is clearly not going to be easy. The pragmatic arguments in favour of teaching statistics (its uses in business, engineering, medicine, agriculture, etc.) may be more persuasive than the ideological arguments ("mathematics for all"). Ultimately, one may expect the same kind of evolution to occur which is affecting mathematics education in the West, but it may have to await larger-scale social and economic changes.

As far as the IASE is concerned, the problems of promoting statistical education are different in degree rather than in kind to those in Western countries. Individual teachers, even at university level, are likely to be more isolated, and face even greater difficulties and prejudices than their Western colleagues. At school level, the problems of influencing the curriculum and tackling the massive problems of teacher training and retraining are rendered just that much more difficult because of the overall economic fragility and social novelty of any kind of mass education programme.

The first task of the IASE, therefore, should be to try and provide professional and moral support for teachers in a basically difficult environment. The Interim Executive, with the help of Anne Hawkins and Lionel Pereira-Mendoza, recently put forward a funding application to UNESCO to help set up a network between teachers in such situations. This initiative should be maintained even if application is not immediately successful. The newsletter and other material could also be of assistance.

Dues may be a problem because of difficulties with the exchange rate and non convertible or only partially convertible currencies. At present the IASE has reduced subscription to members from developing counties, but even so the cost to an individual compared to the benefits obtained may deter many potential members in such countries. Other possibilities, such as institutional membership for university departments, may be worth exploring. The ISI in the past has investigated a number of schemes for assisting statisticians facing such problems. Typically, these involve paying dues in local currency, collecting them locally, and then using them to cover local expenses for international meetings which can be held in their countries, a publications which can be printed there.

The advantages of holding international meetings in such developing countries should be emphasised. It may be the only way a substantial number of teachers, at whatever level, will have the opportunity of participating directly in an international programme, and the experience clearly has a great potential for encouraging local development and emancipation from past problems.

Publications are more difficult but at least the IASE Executive could be alert to the possibilities of saving dollar funds by printing material outside the Western countries with locally generated subscriptions.

6. A summary of suggestions

The following is a summary of points made in the text, which the Association may like to consider as suggestions for possible future actions and activities. Of course, they represent only a preliminary list and are not meant to inhibit the exploration of other possibilities.

1. The Association should discuss with the ISI the extent to which it is expected to take responsibility for the ISI's own objectives in regard to statistical education, and how those might best be discharged.

2. The Association should consider the possibilities of imitating joint activities or projects with the other ISI sections, for example a cycle of joint meetings.

3. Such points may need to be clarified as a basis for future negotiations with bodies such as UNESCO and the Indian government, which currently provide educational funding to ISI.

4. Representatives of specific international organisations could be invited to have observer status on the IASE executive.

5. The IASE Executive should consider setting up a Task Force to solicit and review suggestions for assisting with statistical education in the developing countries.

6. The Association should approach the International Statistical Education Centres to examine the possibilities of assisting them and their staff in their professional roles.

7. The IASE should keep under view the possibility of participating in direct projects to provide resource material, include cheap texts, for teachers and educational groups in developing countries.

8. The IASE should seek opportunities for increasing its membership from developing countries, including institutional membership for university and other teaching departments, and should investigate special arrangements for payment of dues in local currencies.
9. IASE should also consider the opportunities for promoting statistical education in developing countries by holding international meetings there when suitable opportunities arise.
10. The idea of setting up a network for teachers of statistics in developing countries should be endorsed and explored further.

Acknowledgements

I would like to acknowledge the support I have received, during the period I have associated with ISI Education Committee and the IASE, of the ISI Permanent Office staff, the Committee members past and present, and my colleagues at Victoria University, who over this whole period have been extraordinarly tolerant of absences and inroads into their secretarial services.

Bibliography

Barnett V. (ed.) (1982), *Teaching Statistics in Schools throughout the World*, International Statistical Institute, Voorburg.

Bryson K.R. (1991), Using video-based projects in developing countries, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 2, pp. 562-569.

Doyle B. and Brown G. (1991), Statistical training for the South Pacific Commission, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 2, pp. 533-539.

Gani J. (1979), The role of ISI in the field of statistical education (1949-1979), *International Statistical Review*, 47, pp. 237-244.

Gani J. (1987), New challenges to the ISI Education Committee, in R. Davidson and J. Swift (eds.), *Proceedings of the Second International Conference on Teaching Statistics*, University of Victoria, Victoria, Canada, pp. 452-460.

Goudswaard G. (1964), A reappraisal of the statistical education programme of the International Statistical Institute, *Review ISI*, 32, pp. 88-103.

Jain D. (1991), Statistics and Society, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 2, pp. 3-8.

Loynes R.M. (ed.) (1987), *The Training of Statisticians Around the World*, International Statistical Institute, Voorburg.

Nixon J.W. (1960), *A History of the International Statistical Institute, 1885-1960*, International Statistical Institute, The Hague.

Rice S.A. (1948), *An international programme for education in statistics*, UN Document E/CN/3/43, quoted in Nixon (1960).

Rice S.A. (1949), Furtherance of Statistical Education. A provisional programme of the International Statistical Institute, *Review ISI*, 17, pp. 1-34.

Tulya-Muhika S. (1991), Teaching statistics for future government statistical service in Africa, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 2, pp. 552-561.

Vere-Jones D. (1993), *The Coming of Age of Statistical Education* (submitted for publication).

Zarkovich S.S. (1976), A reappraisal of the ISI statistical education programme, *International Statistical Review*, 44, pp. 282-295.

UPDATING TEACHING METHODS IN PROBABILITY AND STATISTICS

Session devoted to Italian school teachers

A report by Maria Gabriella Ottaviani

UPDATING TEACHING METHODS IN PROBABILITY AND STATISTICS

Maria Gabriella Ottaviani
Dipartimento di Statistica, Probabilità e Statistiche Applicate
Università di Roma "La Sapienza", P.le A. Moro 5, 00185 Roma, Italy

The session was attended by 22 teachers, mainly from secondary schools (17). The majority of the participants came from technical schools (ITC), however, teachers from secondary schools specializing in scientific studies (LS) and from industrial secondary schools (ITI) were also present.

It was to be expected that the participants came from these kinds of schools, owing to the fact that the current school legislation provides for "Probability and Statistics" to be on ITC and ITI syllabuses, however, this subject is virtually non-existent and limited to probability only in LSs, with the exception of those LSs where experimental courses have been set up, which follow the new syllabuses of the so called "Brocca Commission".

The participants came mainly from the centre of Italy, where the meeting was held (14), but also from northern and southern Italy (5 and 3 respectively). The session was also attended by a teacher from the Palermo IRRSAE (regional institute for research, experimentation and updating in education).

Upon opening the meeting, the Chair remarked that in order to be governed, an increasingly complex modern society needs increasingly complex statistical systems of information. This calls for, on the one hand, the training of experts able to contribute to the development of the systems, while on the other it demands awareness on the part of citizens, who should be able to use and interpret correctly the data they are offered. This is one of the reasons why "Probability and Statistics" is gaining ground within school syllabuses, from primary to middle and secondary schools. In order for the widening of syllabuses to become effective, however, it is first of all necessary for teachers to be adequately prepared to face their new tasks. Although no specific national survey is available, it seems reasonable to think that currently in Italy "Probability and Statistics" is taught in pre-university schools by teachers who may have studied probability and mathematical statistics, but who almost certainly did not study statistics at university and who consequently find themselves

BEST COPY AVAILABLE

having to tackle the subject practically on their own.

It was also to help teachers in this up-hill work that CIRDIS (Inter-university Research Centre for the Teaching of Statistics) was set up in 1991, mainly on the initiative of the Italian Statistical Society.

For IASE and CIRDIS, this session represented a focal point to establish a link between school and university.

At the moment, some topics, which are both practical and pragmatic, appear of immediate relevance.

Given that teachers have had to face and are still facing a number of difficulties in the teaching of "Probability and Statistics", a few questions come up regarding how teachers relate to the teaching of "Probability and Statistics" in schools, how they think their professional competence could be adequately developed for the future, and what links CIRDIS might be expected to provide in order to meet the needs of pre-university school teachers.

The contributions to the meeting were numerous, varied and constructive. Statistics was generally recognized as a subject playing a formative role from a cultural, social and political point of view and as basically having an interdisciplinary character (Di Pisa, IRRSAE, Palermo). It was also pointed out that, where it had been possible to direct students to carry out small-scale investigations, statistics showed its educational value by improving the students ability to analyze real-life situations and to manipulate data, and it also provided an opportunity for an immediate link with the computer science laboratory (Vitangeli, ITC, Rome).

Despite the unquestionable usefulness and validity of probability and statistics, however, their actual teaching meets with some resistance.

It was clearly pointed out that neither subject was included in the teachers' university courses of study, and that statistics appears to be more orientated towards research, including experimental research, a field with which pre-university education is only marginally if not at all concerned. Finally, it was emphasized that statistics is not an organic part of the Italian school curriculum (Di Pisa, IRRSAE, Palermo). In particular, it was noted that while on the one hand no syllabus coordination exists between the different school levels - primary school, middle school, the first two years and the second three years of secondary school - (Di Pisa, IRRSAE, Palermo), on the other hand, where "Probability and Statistics" is included in the syllabus, i.e. in technical secondary schools, some factors impair its correct teaching, namely its inclusion in the mathematics courses, the non-clear relationship between probability and statistics and the excessive number of topics dealt with (Baldoni, ITI, Gubbio).

In fact, the contents of the courses should be selected in relation to the teaching aims, and in the final evaluation one should have the courage to admit which aims have been achieved and which have not, and then draw the appropriate conclusions (Vitangeli, ITC, Rome).

In connection with this, intervention was advocated at the administrative level as regards the drawing up of syllabuses, this is because lately syllabuses have been broadened, while the hours devoted to mathematics have been reduced. A consequence of this is that if towards the end of the school year it becomes necessary to reduce the planned programme, the cuts will fall on "Probability and Statistics" (Baldoni, ITI, Gubbio). The very interdisciplinary nature of statistics, in fact, reveals itself to be a further obstacle to the teaching of the subject, firstly because of the difficulty of putting together suitable material for the students to work on, then because data referring to different subjects have to be handled in different ways, and finally because it is necessary to have experts who can help reach a satisfactory interpretation of the results (Vitangeli, ITC, Rome). In light of these difficulties, it becomes essential to motivate teachers to widen their knowledge of "Probability and Statistics". Although teachers can easily be culturally motivated, the amount of work required for an accurate introduction of the two subjects, however (De Poli, ITI Venezia-Mestre), also calls for an economic reward for the research carried out (Baldoni, ITI, Gubbio).

The development of the teaching of "Probability and Statistics" through in-service training is now a commonly recognized practice in Italy, but unfortunately it is often fragmentary and improvised.

For this reason, for example, an experimental teaching laboratory for "Probability and Statistics" is being set up in Sicily by IRRSAE, to which mathematics teachers of the various teaching levels, from primary school to university, will give their joint contribution, with the aim of bridging the gap currently existing between secondary school and the lower school levels (Di Pisa, IRRSAE, Palermo).

Research carried out by the Umbrian IRRSAE by means of a questionnaire aimed at teachers, revealed a demand for training in "Probability and Statistics", which was met by setting up a 20-hour course attended by about 60 teachers. However, the impact of this course at student level was not verified (Baldoni, ITI, Gubbio).

Some of participants in the session reported that work packages developed in connection with the training courses aimed at teacher trainers should be available from the Education Ministry (Scotti, LS, Ferrara). It was also observed that the three most recent courses to be held made use of the distance-training system. The systems developed by the Cosenza CUD (Consorzio Università a Distanza) are worth mentioning because of the particular attention devoted to the teaching of "Probability and

Statistics" (Sernini, Padova). Most of the participants, however, were not informed about the existence of this material or about how to obtain it or simply how to have access to it.

The fact that teachers lack information about the various activities in progress and about their results suggests that one of the main objectives could be the co-ordination of all the information concerning the teaching of "Probability and Statistics" carried out by pre-university teachers.

Therefore, CIRDIS could represent a precious link-up opportunity for teachers, by allowing them to exchange ideas, and it could also take up a guiding role by establishing research topics and guidelines (Baldoni, ITI, Gubbio). Morever, the journal "Induzioni - Demografia, probabilità, statistica a scuola", mainly devoted to the teaching of probability and statistics at school, could provide space for the circulation of information about the teaching and teacher-training activities which have been or are still being carried out in Italy.

Although it should be borne in mind that the teachers who attended the meeting are trainers of the Ministry of Education and therefore represent the cream of Italian pre-university teaching, it is however to be noted that the vitality of the debate revealed a real interest in the teaching of probability and statistics, an interest which involves both teachers and students when the usefulness of the two subjects in real-life situations is put to the test.

This interest, however, is held back, on the one hand, by the difficulties of information exchange among teachers and, on the other, by the fact that at present the efforts made and the initiatives developed are not adequately recognized.

POSTER SESSION

Abstracts

Authors:

- Anna Bartkowiak
- Maria Davì
- Yasar Ersoy *et al.*
- Rosa Giaimo *et al.*
- Saleha Naghmi Habibullah
- Ruben Hernandez
- Aldo Mascelloni
- Gianfranco Moncecchi *et al.*
- John Panaretos
- Paola Ranzani *et al.*
- John Taffe
- Evdokia Xekalaki
- Jingyi Yang *et al.*

DRAWING LSE REGRESSION LINE: AN EDUCATIONAL COMPUTER PROGRAM FOR IBM PC

Anna Bartkowiak
Institute of Computer Science, University of Wroclaw
Przesmyckiego 20, 51-151 Wroclaw, Poland

Although some people do not realize it yet fully, educational packages are something different from professional computational packages. Educational statistical packages have as their aim the explanation and reinforcement of some of the concepts and basic ideas of probability and mathematical statistics. When working with the package the student must be active, i.e. being faced with some questions he is forced to make some decisions or to undertake some actions.

PS-STAT elaborated in Wroclaw is such a package for IBM PC computers with EGAVGA or HERC graphical cards. The package is designed as an auxiliary tool for a laboratory supplementing a course for non-mathematical students, e.g. psychologists or biologists. It comprises at present 18 separate modules - each of them can be run separately without keeping the others in the computer core.

One of these modules has as its aim the training of students to draw a regression line. First the data are read in or the rule of generating pseudorandom data is established. The data points (x, y) appear on the screen together with two pivotal points. Each of these pivotal points can be moved (displaced) on the screen so long as the student thinks that a line passing through these two pivotal points would be the LSE regression line. After pressing the F3 and the F4 keys, the student's line and the true LSE line appear (or disappear) on the screen. The respective residual sums of squares and the regression equations are exhibited too. The student can obtain additional information (help) about the principles of estimation the LSE line and the formulae used. Formally the program is part of the package PS-STAT and can be called from a common menu, none the less it can be run separately without keeping the others in core.

The package was partially sponsored by the Polish Research Council and can be obtained free after sending two HD diskettes to the authors.

NEW OPENINGS FOR TEACHING APPLIED STATISTICS IN THE ACADEMIC COURSES

Maria Davì
Istituto di Statistica Applicata alla Ricerca Economica
Facoltà di Economia e Commercio
Viale delle Scienze, 90128, Palermo, Italy

This poster shows the results of a survey carried out in order to verify the diffusion of the teaching of applied statistics at the Faculties of Economics and Commerce and Statistical Sciences in the Italian University.

Particular attention has been paid to the identification of disciplines that, in recent years, have increasingly required the adoption of statistical methodology as an analytical tool.

The research has been extended to the fields of interest for the applicability of technical methods of statistics. Consideration has also been given to the developments achieved in those subjects which have existed for a long time in the Italian university didactic system (demography, social statistics and economic statistics).

Information has been obtained through a questionnaire sent by mail and submitted to instructors of applied statistics who are also members of the Italian Statistical Society.

The questions were fundamentally concerned with the framework of the courses with regard to possible collaborators, to the features of the statistical work done with students and also to new teaching methods adopted, if any.

Instructors have been invited to express their opinions in relation to the connections between their subject and the disciplines which are propaedeutical to it, and those ones with respect to which applied statistics has a supportive role.

Moreover, some remarkable shortcomings came out during the interview due to the lack of co-operation and co-ordination between statistical subjects and the disciplines with probabilistic-mathematical contents.

Instructors also complained about the limited availability of stable collaborators and about the lack of equipment suitable for the more sophisticated methods.

Nevertheless, the increase and the great range of today's methods used in statistical training are exhibited together with the traditional ones, as both constitute a substantial support to the teaching of applied statistics.

Finally, some proposals are outlined with regard to the efficient reorganization of academic courses, on the basis of indications which arose from the survey.

STUDENTS' VIEWS ON EDUCATIONAL ASPECTS OF STOCHASTICS

Yasar Ersoy and Tibor Nemetz
Middle East Technical University, 06531 Ankara, Turkey

This presentation summarizes answers of 250 students of the Middle East Technical University who are subjected to a compulsory course in probability and/or statistics. A questionnaire, which aimed at collecting information about:

- their views on the educational aspects of statistics;
- their personal views and experience of other people's perception of the random phenomena and ways of thinking;

was administrated in various groups of university students either before or after taking such a course in January 1993.

The poster detailes the questions asked, and a number of bivariate contingency tables and the frequency distributions of some items.

We are aware that our data provides just a rough orientation. Nevertheless, some of the findings may be generally valid. We have not found any significant difference between the answers of boys or girls. Neither does there appear to be any significant difference between the age groups or students of different attendance years. It is somewhat surprising that students' opinions do not depend strongly on the number of courses completed. The difference can be spotted, however, if the courses were completed with good grades.

It is also instructive that more than 40 per cent of the students has definite opinions on how to improve teaching. The majority of them consider everyday examples, classroom experiments and discussions of games as important means in this direction. Some of them urge an improvement of the quality of teacher training or suggest increasing the allocated time.

THE TEACHING OF STATISTICS AT THE UNIVERSITY LEVEL

Rosa Giaimo and Anna Maria Milito
Istituto di Statistica, Facoltà di Economia e Commercio
Viale delle Scienze, 90128 Palermo, Italy

The recent inclusion of Statistics in faculties where it is of an applied nature raises questions concerning the diverse contents and approaches in the various fields of study and also the instructors' adaptability in teaching the peculiar features of each field.

To determine how statistics courses are inserted in the different university degree programmes we sent out a questionnaire with a few, essential questions about the instructor's position, classes offered, textbooks and possible experimental didactic methods.

On a first glance basis, we have observed that the instructors generally are on the roll of other faculties. We have also noted that the contents and textbooks per course differ greatly and generally the contents are taught using traditional methods.

The responses appear to show that close collaboration between statisticians and the various field instructors has not yet been established.

STATISTICAL EDUCATION IN MEXICO

Ruben Hernandez
Instituto Tecnologico Autonomo de Mexico
Rio Hondo 1, San Angel, C.P.01000 Mexico

The teaching of statistics at university level in Mexico is a relatively recent development. The majority of statisticians with doctorates have completed their studies in other countries. This study offers an analysis of the situation as regards the teaching of statistics at university. In this case emphasis is placed primarily on two examples of the National Autonomous University of Mexico (UNAM) and the Mexican Autonomous Institute of Technology (ITAM). Two different aspects are considered: the professions that produce statisticians and those which make use of statistics as an aid to their work.

REACHING OUT TO INSTITUTIONS BEYOND LAHORE

Saleha Naghmi Habibullah
Kinnaird College for Women, Lahore, Pakistan

One of the main problems faced by the existing system of statistical education in Pakistan is that, during the past few years, the programs of study have not been application oriented. Undue emphasis on textbook material at the FA/FSc and BA/BSc levels promotes in the students rote-learning of definitions and formulae, and impedes students' ability to apply statistical concept and techniques to real-world situations.

The academic year 1991-92 witnessed a number of efforts made by the author to promote the idea of project-based teaching and learning of introductory statistics in the city of Lahore. These included Kinnaird's Visiting Statistical Exhibition (KVSE) which aimed at motivating students of various colleges to participate in Kinnaird's Inter-collegiate Statistical Competition (KISC), and was taken to thirteen different colleges of Lahore.

The academic year 1992-93 has witnessed efforts aimed at extending KVSE and KISC to other cities and towns of Pakistan. In doing so, the objective is to spread the message to various parts of the country that small-scale statistical projects can prove to be very interesting for students, not to mention the benefits of such projects in enhancing students' understanding of statistical concepts and techniques. Wide-spread acceptance of statistical project-work at the FA/FSc and BA/BSc levels may play a key role in the improvement of statistical education in Pakistan.

This paper presents the methodology that was adopted to extend KVSE and KISC to other towns/cities during 1992-93, the response received from the institutions that were approached, and some possible strategies for greater success in future years.

SPRINT - ANALYSIS OF A GAME

Aldo Mascelloni
Scuola Media Statale Giovanni Pascoli, Grosseto, Italy

Sprint is a simple game for children: there are four runners (the Hare, the Squirrel, the Tortoise and the Snail) each one with a given speed (expressed by the values shown on the sides of the relative dice, that is, 4-3-3-2-2-2, 3-3-2-2-1-1, 2-2-2-1-1-0, 1-1-1-1-0-0, respectively), who must complete four different rounds of the length, in the order, of 32, 24, 16 and 8 squares. At each turn, each runner throws his dice and goes forward the number of steps indicated.

To win two different hypotheses were taken into consideration:

a) The person who reaches the last square first wins (without needing to pass it);

b) the person who passes the last square wins, and if two players pass the last square at the same time, the one who goes furtherest wins.

A class experience (in a third year elementary) resulted in the following conclusions:

Under hypothesis a) each runner needed an average of 12 throws to finish the round. Therefore equal probability of winning; the game was said to be well balanced.

Under hypothesis b) it was maintained that the Hare had an advantage, followed by the Squirrel, the Tortoise and the Snail, in that order.

But computer simulation using a Pascal programme gave different results:

Under hypothesis a) the winning percentages out of 80,000 races were: Hare 16.06%, Squirrel 20.54%, Tortoise 27.16%, Snail 36.24%.

Under hypothesis b) the winning percentages out of 80,000 races were: Hare 22.97%, Squirrel 24.8%, Tortoise 27%, Snail 25.16%.

The most important aspect from a didactic point of view isn't to have found the solution of the game, but to have discovered what the computer can do; it allowed us to notice the error and helped us to remediate it.

TEXTBOOKS AND STATISTICS IN ITALIAN PRIMARY SCHOOL

Gianfranco Monceochi and Maria Pia Perelli D'Argenzio
Dipartimento di Matematica, Università Cattolica
Via Trieste 17, 25121 Brescia, Italy

The poster presents a critical comparison between the objectives of the educational material covered under probability and statistics for the primary school and some of the teaching proposals presented in textbooks and teacher's guides.

In the statistical section there is a critical analysis of the most common approaches:

- It is noted that, in general, no difference is made between quantitive characters and qualitative ones.
- Errors in graph construction are observed.
- Reference is made to the fact that there is a constant incoherence between the minimum and very important objectives posed by Education Ministry Programmes and the teaching proposals that tend to convey a much wider array of subjects but in a mistaken, confused and superficial manner.

In the probability section there is an indication of the characteristics that are generally found in textbooks and teachers' guides:

- Some of the proposals introduce the classical measurement of probability already in the first year or at the beginning of the second year of the primary level, and then move on to explain situations that apply in complex events. This material is not requested in the programmes and is very difficult to understand.
- These proposals consider the terms "probability" and "possibility" as synonimous; the term "probable" is used as synonimous of "easy" or "frequent".
- Very often the classical probability is introduced by means of frequency methods that are based on coin tossing or dice rolling, both of which exclude the fundamental aspects of uncertain situation observance (information analysis) and substitutes it with some trials.
- The probability value found by means of frequency seems to correspond to the trials already done; no importance is given to the

assumption that this can be taken as measurement of probability that results from random experiments.

- The number of trials is decided in an arbitrary way.
- The probability measurement using frequencies, which is relevant in cases where classical measurement cannot be applied, is totally ignored.

USES AND MISUSES OF STATISTICAL METHODS IN FORMULATING AND IMPLEMENTING EDUCATIONAL POLICIES IN A CENTRALLY CONTROLLED EDUCATIONAL SYSTEM

John Panaretos
Department of Statistics
Athens University of Economics and Business
76, Patision St., 10434 Athens, Greece

The Greek government that took office in the Spring of 1990 decided to make changes at all levels of education in Greece.

In order to make the intended changes more easily acceptable to voters the Ministry of Education decided to send out questionnaires to all persons that could be affected (directly or indirectly) by the intended changes.

As far as university education was concerned questionnaires were sent to all faculty members of all the universities in Greece, to university students (through their representatives) and to interested parents.

For primary and secondary education similar questionnaires were sent to teachers (through their unions) to parent groups and to students.

The negative aspects of the results of the approaches and the methodology employed, as well as the misuse of statistical methods by the government in order to achieve its political objectives are discussed.

STUDY OF DISTRIBUTIONS OF SAMPLE MEAN AND SAMPLE VARIANCE OF RANDOM SAMPLES FROM A STANDARD NORMAL DISTRIBUTION: COMPUTING SIMULATION

Paola Ranzani and Morena De Poli
I.T.I.S. C. Zuccante, 30174 Venezia-Mestre, Italy

This experience, made in a statistical laboratory for the more advanced class in a technical institute for computer science, is realized to verify that the sample mean and the sample variance have, respectively, a normal distribution and a chi-square distribution.

The steps to follow are:

1) to generate m random samples each of n elements from a standard normal distribution;
2) to calculate, for each sample, the mean and the sum of the squares of the observations;
3) to make the frequency distribution of the m sample means and of the m sample variances (grouping the observations into k class intervals);
4) to determine, for each distribution, the probability, p_i, that the particular hypothesized distribution would assign to the *ith* class interval by a numerical method of integration (Bezout, Cavalieri Simpson), and to calculate the expected number, $m \times p_i$, of observations in the *ith* class interval ($i = 1,2,...,k$);
5) to apply the chi-square test to verify the goodness of fit.

The result of the chi-square test and the graphic comparison between the generated distributions and the hypothesized distributions by modifying the input parameter m and n allow to confirm what it was studied during the theoretical lessons.

MULTI-MODAL LEARNING IN STATISTICS

John Taffe
Mathematics Department, Swinburne University of Technology
Hawthorn, 3122 Australia

A pilot project in ***multi-modal learning*** (MML) is under way this year for Applied Science students at Swinburne's small Eastern Campus (EC). The introductory applied statistics unit "Data analysis and Probability" is one of four taught by this method. MML involves students in different "modes" of interaction with staff and peers on the main campus (i.e., EC). Other modes involve the student working at home and communicating via e-mail and electronic bulletin board with staff and peers (home mode) and meeting at local "learning centres" for remote tutorials and discussion groups (sub-core mode).

A first requirement of teaching in a MML context is the preparation of a comprehensive "Learning Guide".

The poster discusses progress so far with MML as a whole and the Learning Guide.

STATISTICAL EDUCATION IN GREECE

Evdokia Xekalaki
Department of Statistics
Athens University of Economics and Business
76, Patision St., 10434 Athens, Greece

Statistical education has not made great steps at the primary and secondary levels of education in Greece. As a result, school leavers who seek entrance to the university do not have a very clear idea as to what exactly statistics is. On the other hand, potential employers of university graduates both in the public as well as in the private sector appear not to be fully aware of the advantages of employing persons with statistical training.

The opportunities and constraints faced in an environment of such factors, among others, by the Department of Statistics of the Athens University of Economics and Business are discussed, four years after its establishment.

CURRENT DEVELOPMENT OF STATISTICAL EDUCATION IN CHINESE AGRICULTURAL UNIVERSITIES

Jingyi Yang and Nan Li
Jilin Agricultural University, Changchum, 130118 Jilin, P.R. China

1. A comparison of statistical education contents

In the early 60s, statistical education in Chinese agricultural universities was very limited. Four year B.Sc students had only about 20 hours of statistical education and 50 hours of mathematics on the average. Only a few staff could give lectures on the level of multiple regression in these universities.

There were no significant changes until 1976. However, a big innovation occurred in 1977. During the 1980s, calculators were widely used in classes. The lecture programs in statistics increased to 60-70 hours plus 20-30 hours practice. Probability and statistics, linear algebra and computer programming were designed for first degree students. However, there was a slower increase in statistical education programs in the 1990s due to the limitation of computer facilities.

Table 1. *A list of statistical education contents and lecture hours on B.Sc students (Jilin Agricultural University, P. R. China)*

	1960s	1980s	1990s
Mathematics	50	90	120
Linear algebra	0	20	30
Probability & statistics	0	30	40
Biometry	20	80	90
Sample statistics	12	10	10
Experimental design	2	10	20
Analysis of variance	4	25	20
Regression	2	35	40

A survey example of Jilin Agricultural University, P.R. China is given in Table 1. The results, to some extent, represent the average level of statistical education in about 80 Chinese agricultural and forestry universities. It has reached the theoretical level of the current British education on M.Sc. programs of non biometry students.

2. Problems and discussion

(1) Computer facilities in Chinese agricultural universities are much more limited in both numbers and efficiency. The lack of computers is caused by financial problems. The lower efficiency of computers is due to software problems and computer supported services systems.

(2) There is a gap between those who teach mathematical statistics and those who teach biometry. Statistical instructors lack knowledge in biological science, and on the other hand, most biometry teachers are agronomists who are short of statistical knowledge. This gap will be a big obstacle to the further development of statistical training and education in agricultural universities.

(3) Existing teaching materials, such as text books and teaching facilities, need to be enriched by current statistical theories and methods as the majority of text books in Chinese agricultural universities were edited and written in early 1980s. Some of the contents are out of date and simplistic, and new statistical methods are not included.

(4) Statistical service centres and computer service centres need to be established in agricultural universities so as to meet the need of further development of statistical education in agricultural universities.

LIST OF PARTICIPANTS

LIST OF PARTICIPANTS

Afsarinejad Kasra
Dept. of Mathematics
Linköping University
S-58183 Linköping SWEDEN
tel.: +46 13282198
fax: +46 13100746
E-mail: Kaafs@math.liu.se

Ambroise Daniel
Université Pierre et Marie Curie
4 Place Gossier
75252 Paris Cedex 05 FRANCE
tel.: +33 44273323

Andersen Otto
Danmarks Statistik
Sejrögade 11
02100 Copenhagenö DENMARK
tel.: +45 39173231
fax: +45 31184801

Anton Peter
Dept. of Mathemathical Statistics
University of Umeå
S-90187 Umeå SWEDEN
tel.: +46 90166399
fax: +46 90167658
E-mail: spwa@matstat.umu.se

Aprea Marcello
I.T.I.S. Galilei
Via Galilei 68
57100 Livorno ITALY
tel.: +39 586 404394
fax: +39 586 429548

Arnold Tim
Dept. of Statistics
Instructional Computing NCSU
Raleigh, NC 27695 U.S.A.
tel.: +1 919 5151927
fax: +1 919 5157591
E-mail: arnold@stat.ncsu.edu

Badoer Giovanna
I.T.C. "Calvi"
Via Palermo 5
35142 Padova ITALY
fax: +39 49 650435

Baldoni Corrado
I.T.I.S
Via Bottagnone
06024 Gubbio (PG) ITALY
tel.: +39 75 9277641

Bangdiwala Shrikant I.
Dept. of Biostatics, CB#7400
University of North Carolina
Chapel Hill
NC 27599-7400 U.S.A.
tel.: +1 919 962 3266
fax: +1 919 962 3265
E-mail: kant@uncmvs.oit.unc.edu

Barnett Vic
Rothamsted experimental station
Harpenden, Herts, AL5 2JQ U.K.
tel.: +44 582 467116
fax: +44 582 467116
E-mail: Vbarnett@resa.afrc.ac.uk

Bartkowiak Anna
Institute of Computer Science
University of Wroclav
Przesmyckiego 20
51-151 Wroclaw POLAND
tel.: +48 71 251271
E-mail: ABA@plwruw11.bitnet

Bécue Mónica
Dept. Estadistica
Facultad d'Informàtica
Universitat Politecnica de Catalunya
Pau Gargallo 5
08028 Barcelona SPAIN
tel.: +34 3 401 7000
fax: +34 3 401 7040
E-mail: Monica@eio.upc.es

Beffa Silvana
S. Angelo di Celle
06050 Deruta (PG) ITALY
tel.: +39 75 974340

Belli Gabriella
VPI & SU
2990 Telestar Court
Falls Church, VA 22042 U.S.A.
tel.: +1 703 6986064
fax: +1 703 6986068
E-mail: belli@vtvm1.bitnet
or belli @vtvm1.cc.vt.edu

Biehler Rolf
Inst. für Didaktik der Math.
IDM Universitat
Postf. 100131
D - 33501 Bielefeld GERMANY
tel.: +49 521 106 5058
fax: +49 521 106 2941
E-mail: rolf.biehler@post.uni-bielefeld.de

Blacona Maria Teresa
Facultad de Ciencias Economicas
Universidad de Rosario
Boulevard Oroño 1261
02000 Rosario ARGENTINA
tel.: +54 41 4030
fax: +54 41 832388

Blejec Andrej
Institute of Biology
University of Ljubljana
Karlovska 19, P.O. box 141
61001 Ljubljana SLOVENIA
tel.: +38 61 211744
fax: +38 61 211735
E-mail: andrej.blejec@uni-lj.si

Blumberg Carol J.
Dept. of Math and Stat
Winona State University
Winona, MN 55987-5838 U.S.A.
tel.: +1 507 4575589 4575370
fax: +1 507 4575681
E-mail:
wncarolj@vax2.winona.msus.edu

Borovcnik Manfred G.
Inst. für Mathematik und Statistik
Universität Klagenfurt
Sterneckstrasse 15
9020 Klagenfurt AUSTRIA
tel.: +43 463 2700778
fax: +43 463 2700759
E-mail: borovcnik@edvz.uni-klagenfurt.ada.at

Borsini Ido
ITGC Corridoni
Via B. Bambozzi 8
60027 Osimo (AN) ITALY
tel.: +39 71 78 19572
fax: +39 71 71 33478

Box George
Center for Quality & Productivity
Improv., University of Wisconsin
610 Walnut Street
Madison, WI 53705 USA
tel.: +1 608 263 2520
fax: +1 608 835 9077

Bracalente Bruno
Dip. di Scienze Statistiche
Università di Perugia
Via A. Pascoli
06100 Perugia ITALY
tel.: +39 75 5855242
fax: +39 75 43242

Braumann Carlos A.
Dept. de Matemática
Universidade de Évora
P-7000 Évora PORTUGAL
tel.: +351 66 26201
fax: +351 66 20775

Brunelli Lina
Dip. di Scienze Statistiche
Via A. Pascoli
06100 Perugia ITALY
tel.: +39 75 5855242
fax: +39 75 43242
E-mail: statli@ipguniv.bitnet

Bryson Kenneth R.
Jecor Unit 61306
Box 85
Apo, AE 09803-1306 U.S.A.
tel.: +1 966 1 464 0433
fax: +1 966 1 464 5032

Camillo Furio
Dip. di Scienze Statistiche
Università di Bologna
Via Belle Arti 41
40126 Bologna ITALY
tel.: +39 51 258186
fax: +39 51 232153

Capiluppi Claudio
Dip. di Scienze Statistiche
Università di Padova
Via S. Francesco 33
35121 Padova ITALY
tel.: +39 49 832140
fax: +39 49 8753930
E-mail: capilup@ipdunivx.unipd.it

Caredda Carla
Dip. di Matematica
Università di Cagliari
Viale Merello 92
09123 Cagliari ITALY
tel.: +39 70 2000416
fax: +39 70 2000420

Carriòn Garcia Andrés
Dep of Statistics,
Universidad Politecnica de Valencia
Cno. de Vera s/n
46020 Valencia SPAIN
tel.: 34 6 3877493
fax: 34 6 3877499

Castles Ian
Australian Bureau of Statistics
P.O. Box 10,
Belconnen Act 2616 AUSTRALIA
tel.: +61 6 2526705
fax: +61 6 2531328

Chaseling Janet
School of Applied Maths & Statistics
Division of Environmental Sciences
Griffith University
Queensland 4111 AUSTRALIA
tel.: +61 7 875 7515
fax: +61 7 875 7459
E-mail: J.Chaseling@ens.gu.edu.au

Cicchitelli Giuseppe
Dip. di Scienze Statistiche
Università di Perugia
Via A. Pascoli C.P. 1315 - Succ. 1
06100 Perugia ITALY
tel.: +39 75 5855242
fax: +39 75 43242
E-mail: stat7@ipguniv.bitnet

Clark Megan
Math. Educational Unit & Inst. of
Stat. and Operations Research
Victoria University
P.O. Box 600
Wellington NEW ZEALAND
tel.: 64 4 4721000
fax: 64 4 4955118
E-mail: dvj@isor.vuw.ac.nz

Cohen Steve
Curricular Software Studio
Arena Computer Annex
Tufts University
Medford, Massachusetts 02155 USA
tel.: +1 617 627 3004
fax: +1 617 627 3667
E-mail: scohen@pearl.tufts.edu

Colagrande Vittorio
Via Vito da Ortona 5
66026 Ortona (CH) ITALY

Conrad Simon
UMIST
P.O. Box 83
Manchester M60 1Q U.K.
tel.: +44 61 2003429

Cortinovis Ivan
Ist. Statistica Medica e Biometria
Via G. Venezian, 1
20133 Milano ITALY
tel.: +39 2 23 61 302
fax: +39 2 23 62 930
E-mail: pgduca@imi.cilea

Cruciani Lina
Istituto Tecnico Commerciale
L. Einaudi
63018 P.to S. Elpidio (AP) ITALY

Daddi Pierluigi
Dip. di Scienze Statistiche
Università di Perugia
Via A. Pascoli
06100 Perugia ITALY
tel.: +39 75 5855242
fax: +39 75 43242

Dagnelie Pierre
Statistique et Informatique
Faculté des Sciences agronomiques
Avenue de la Faculté d'Agronomie 8
5030 Gembloux BELGIQUE
tel.: 32 81 62 2478
fax: 32 81 61 4544
E-mail: dagnelie@bgxfsa51

Davì Maria
Ist. Stat. Applic. alla ricerca econ.
Università di Palermo
Viale delle Scienze
90128 Palermo ITALY
tel.: +39 91 6521090
fax: +39 91 6521095

De Alba Enrique
Div. of Mathematics
I.T.A.M.
Rio Hondo 1
San Angel, C.P. 01000 D.F. MEXICO
tel.: +52 5 5486902
fax: +52 5 5486902
E-mail: dealba@itamvms1.bitnet

De Cani John S.
Dep. of Statistics
University of Pennsylvania
3008 Steinberg - Dietrich Hall
Philadelphia, PA 19104-6302 U.S.A.
tel.: +1 215 898 8222
fax: +1 215 898 1280

De Poli Morena
I.T.I.S C. Zuccante
Via Baglioni 22
30173 Venezia - Mestre ITALY

Diamanti Settimio
I.T.C. L. Einaudi
63018 P.to S. Elpidio (AP) ITALY

Di Pisa Giovanni
Liceo Scientifico Basile
90100 Palermo ITALY

Edwards Howard Paul
Dept. of Statistics
Massey University
Private Bag 11222
Palmerston North
NEW ZEALAND
tel.: +64 6 3569099
fax: +64 6 3505611
E-mail: H.Edwards@massey.ac.nz

Eisenbach Ronit
Dept. of Statistics
Tel Aviv University
Ramat Aviv
Tel Aviv 69978 ISRAEL
tel.: +972 2 522455
fax: +972 3 6409357

Ferrandi Maria Angela
Universidad Nacional de Rosario
Herrera 4030
2000 Rosario-P.to S.ta Fe
ARGENTINA
fax: +54 41 43030

Forcina Antonio
Dip. di Scienze Statistiche
Università degli Studi di Perugia
Via A. Pascoli
06100 Perugia ITALY
tel.: +39 75 5855242
fax: +39 75 43242

Frierson Dargan
Dept. of Mathematical Sciences
UNC-Wilmington
601 South College Road
Wilmington, NC 28403-3297 U.S.A.
tel.: +1 919 395 3113
fax: +1 919 3953114
E-mail: frierson@vxc.uncwil.edu

Galmacci Gianfranco
Dip. di Scienze Statistiche
Università degli Studi di Perugia
Via A. Pascoli
06100 Perugia ITALY
E-mail: glm@fisher.stat.unipg.it

Garcia Maria del Carmen
Facultad de Ciencias Economicas
Universidad Nacional De Rosario
Boulevard Oroño 1261
2000 Rosario ARGENTINA
tel.: +54 41 21025
fax: +54 41 43030

Ghilardi Gilberto
Dip. Statistica e Matematica
Università di Pisa
Via Ridolfi 10
56124 Pisa ITALY
tel.: +39 50 549255
fax: +39 50 541437

Giaimo Rosa
Istituto di Statistica
Facoltà di Economia e Commercio
Università di Palermo
V.le delle Scienze
90128 Palermo ITALY
tel.: +39 91 485726
fax: +39 91 485726

Giovagnoli Sandra
Dip. di Matematica ed Informatica
Università di Udine
33100 Udine ITALY

Gunsjö Anna
Dept. of Statistics
Uppsala Universitet
P.O. Box 513
S-751 20 Uppsala SWEDEN
tel.: +46 18 18 11 40
fax: +46 18 55 44 22
E-mail: Anna.Gunsjo@statistik.uu.se

Hanoune Nelly
Groupe ASU Formation
Université des Sciences et
Technologies de Lille
59655 Villeneuve d'Ascq Cedex
FRANCE
tel.: +33 20 43 6760
fax: +33 20 43 6774
E-mail: hanoune@alea.citilille.fr

Hawkins Anne S.
64 Bedford Court Mansions
Bedford Ave, Bloomsbury
London WC1B 3AD U.K.
tel.: +44 71 6360058
fax: +44 71 6360058
E-mail: teexash@ioe.ac.uk

Hernandez Ruben
I.T.A.M.
Apartado Postal 70.429
04510 Coyoacan D.F. MEXICO
tel.: 52 2 280 1488
fax: 52 5 548 6902
E-mail:
Hernandez@musthe:itamvms1.bitnet

Holmes Peter
Centre for Statistical Education
University of Sheffield
Sheffield S3 7RH U.K.
tel.: +44 742 768555
fax: +44 742 824292 ext. 4117
E-mail: p.holmes@pa.shef.ac.uk

Iversen Gudmund
Swarthmore College
500 College Avenue
Swarthmore PA 19081 U.S.A.
tel.: +1 215 328 8514
fax: +1 215 328 8673
E-mail: giverse1@cc.swarthmore.edu

Jackson Fraser
Victoria University of Wellington
P.O. Box 600
Wellington NEW ZEALAND
fax: +64 4 232 4440
E-mail: jackson@matai.vuw.ac.nz

Jolliffe Flavia R.
School of Maths, Stats and Computing
University of Greenwich
Wellington Street, Woolwich
London SE18 6PF U.K.
tel.: +44 81 3168724 or 3168700
fax: 44-81-3168665
E-mail: f.r.jolliffe@uk.ac.gre

Kibirige George W.
Inst. of Statistics & Appied Economics
Makerere University
P.O. Box 7062
Kampala UGANDA
fax: +256 41 245 888

Kilian H.
Institut für Didaktik der Mathem
Universität Dortmund
Vogelpothsweg 87
D-44221 Dortmund GERMANY
tel.: +49 231 755 1
fax: +49 231 755 5307

Konold Clifford
SRRI
Hasbrouk Laboratory
University of Massachusetts
Amherst, MA 01003 U.S.A.
tel.: +1 413 545 5886
fax: +1413 545 1691
E-mail: konold@srri.umass.edu

Kovacsics Katherine
University Eötvös Loránd
Egyetem tér 1-3
Budapest H-1364 HUNGARY
tel.: +36 1 266 5945
fax: +36 1 266 8055

Laurent-Duhamel Marie-Jeanne
1215 Av. Amiral Landrin
64110 Jurançon FRANCE
tel.: +33 5906 0691
fax: +33 5984 1696

Lazzaro Irene
Via E. Galli 13
87100 Cosenza ITALY
tel.: +39 984 33357

Lochert Paul B.
Dept of Mathematics
Monash University - Caulfield
Campus
P.O. Box 197, Caulfield East
Melbourne, Victoria 3145
AUSTRALIA
tel.: +61 3 5732337
fax: +61 3 5732227
E-mail: PLochert@monash.edu.au

Lombardo Enzo
Dip. Studi Geoeconomici
Via del Castro Laurenziano 9
00161 Roma ITALY
tel.: +39 6 49766425
fax: +39 6 4957506

Lorenz R. J.
Federal Res. Cent. for Virus Diseases
of Animals
Postfach 1149
D-72001 Tubingen GERMANY
tel.: +49 7071 603 329
fax: +49 7071 603201

Maiotti M. Lorella
Scuola Elementare
Ferro di Cavallo
06100 Perugia ITALY

Mascelloni Aldo
Scuola Media G. Pascoli
58100 Grosseto ITALY

Mastrodonato Antonio
Istituto di Statistica
Facoltà Economia e Commercio
Università di Bari
Via C. Rosalba 53
70124 Bari ITALY

Mattei Ildo
Ist. Tecnico Benincasa
Via Raffaello Sanzio 66
60035 Iesi (AN) ITALY

McKenzie, Jr. John D.
Babson College
Babson Park, MA 02157-0310 U.S.A.
tel.: +1 617 239 4479
fax: +1 617 239 5232
E-mail: MCKENZIE@Babson.bitnet

Mecozzi Sabrina
I.T.C. L. Einaudi
63018 P.to S. Elpidio (AP) ITALY

Mejza Stanislaw
Dept of Mathematical & Statistical Methods
Agricultural University of Poznan
ul. Wojska Polskiego 28
PL-60-637 Poznan POLAND
tel.: +48 61 224581
fax: +48 61 44125

Meksi Ermelinda
Facoltà di Economia e Commercio
Università di Tirana
Tirana ALBANIA

Milito Anna M.
Ist. Statistica
Facoltà di Economia e Commercio
Università di Palermo
Viale delle Scienze
90128 Palermo ITALY
tel.: +39 91 6563314
fax: +39 91 485726

Militký Jiří
Dept. of Textile Materials
TU Liberec
461 17 Liberec CZECH REP.
tel.: +42 48 21073
fax: +42 48 21073
E-mail: jiri.militky@vslib.cz

Minciaroni Adonella
Scuola Media Statale
M. Buonarroti
Via O. Baldoni 2
06065 Passignano (PG) ITALY

Moncecchi Gianfranco
Seminario Didattico
Università Cattolica di Brescia
Via Trieste, 17
25121 Brescia ITALY

Moore David S.
Department of Statistics
Purdue University
W. Lafayette, IN 47907-1399 U.S.A.
tel.: +1 317 494 6050
fax: +1 317 494 0558
E-mail: dsm@stat.purdue.edu

Morin Anne M. J.
IRISA
Université de Rennes 1
Rennes 35042 FRANCE
tel.: +33 9984 7222
fax: +33 9938 3832
E-mail: amorin@irisa.fr

Moscatelli Antonio
I.P.S.I.A.
06012 Città di Castello (PG) ITALY

Nemetz Tibor
Dept. of Sci. Education
Middle East Techn. University
06531 Ankara TURKEY
fax: +90 4 210 1112
E-mail: nemetz@trmetu.bitnet

BEST COPY AVAILABLE

Nikolov Hristo
Department of Statistics
University of Zimbabwe
P.O. Box MP 167
Harare ZIMBABWE
tel.: +263 4 303211
fax: +263 4 732828
E-mail: nikolov@zimbix.rh.ac.za

Nüesch Peter E.
EPF - MA
CH 1015 Lausanne SWISSE
tel.: +41 21 693 2592
fax: +41 21 693 4250
E-mail: nuesch@elma.epfl.ch

Ottaviani Maria Gabriella
Dip. Stat., Prob. e Stat. Appl.
Univ. degli Studi di Roma "La
Sapienza", P.le A. Moro 5
00185 Rome ITALY
tel.: +39 6 49 58308
fax: +39 6 49 59241

Panaretos John
Dept. of Statistics
Athens University of Economics and
Business
76 Patision St.
Athens 10434 GREECE
tel.: +30 1 8230488
fax: +30 1 8210509

Pannone Maria A.
Dipartimento di Scienze Statistiche
Università degli Studi di Perugia
Via A. Pascoli
06100 Perugia ITALY
tel.: +39 75 5855242
fax: +39 75 43242
E-mail: pnn@ipguniv.bitnet

Passamonti Maria Elena
Universidad Nacional de Rosario
Vera Musica 1118
2000 Rosario ARGENTINA
tel.: +54 41 398093
fax: +54 41 43030

Perelli Maria Pia
Dipartimento di Matematica
Università Cattolica di Brescia
Via Trieste 17
25121 Brescia ITALY
tel.: +39 30 52057
fax: +39 30 51223

Perry Mike
Department of Mathematical Sciences
Appalachian State University
Boone, NC 28608 U.S.A.
tel.: +1 704 2623050
E-mail: perrylm@alf.appstate.edu

Pessoa Djalma
IBGE
Av. Franklin Roosvelt, 166
Rio de Janeiro, RJ BRAZIL
tel.: +55 21 220 6671
fax: +55 21 262 7308
E-mail: DGC@bribge.bitnet

Phillips Brian
Department of Mathematics
Swinburne
University of Technology
P.O. Box 218 Hawthorn
Victoria 3122 AUSTRALIA
tel.: +61 3 8198283
fax: +61 3 8190821
E-mail: brp@stan.xx.swin.oz.au

Polfeldt Thomas
Statistics Sweden
SKOL, Statistics Sweden
S-115 81 Stockholm SWEDEN
tel.: +46 8 783 4760
fax: +46 8 783 4599

Postelnicu Tiberiu
Centrul de Statistica Matematica
Bd. Magheru 22
70 1 58 Bucuresti 22 ROMANIA
tel.: +40 1 50 5755
fax: +40 1 312 0209

Pucci Angelica
Scuola Media Dante Alighieri
Via Scipione Lapi 16/ak
06012 Città di Castello (PG) ITALY
tel.: +39 75 8520871

Puranen Juha
Mannerh. tie 55A24
00250 Helsinki FINLAND
tel.: +358 0 1912538
fax: +358 0 1912124
E-mail: jpuranen@noppa.helsinki.fi

Råde Lenmart
Mathematics Dept.
Chalmers University of Technology
S 41296 Gothenburg SWEDEN
fax: +46 31 161973
E-mail: abk@math.chalmers.se

Raha A.B.
Internat. Statistical Education Centre
203 Barrackpore Trunk Road
Calcutta 700 035 INDIA
fax: +91 33 566680
E-mail: isec@isical.ernet.in

Ranzani Paola
I.T.I.S. C. Zuccante
Via Baglioni 22
30173 Venezia - Mestre ITALY

Regier Mary
19577 Misty Lake Drive
Cleveland Ohio 44136 U.S.A.
tel.: +1 216 5722710
fax: +1 216 3685163
E-mail: mhr@po.cwru.edu

Reiczigel Jenö
Dept. of Biomath.
Univ. of Veterinarisci
István út 2
H-1078 Budapest HUNGARY
tel.: +36 1 2696871
fax: +36 1 2696871
E-mail: H1220rei@ella.hu

Ricci Vincenzo
Istituto Tecnico Commerciale
V. Emanuele II
P.le Anna Frank
06100 Perugia ITALY

Rossi Carla
Dip. di Matematica
Università degli Studi di Roma
"Tor Vergata"
Via Orazio Raimondo
00173 Roma ITALY
tel.: +39 6 72594676
fax: +39 6 72594699
E-mail: rossi@mat.utovrm.it

Rouncefield Mary
Chester College of Higher Education
Cheyney Road
Chester, CHI 4BJ U.K.
tel.: +44 244 375444
fax: +44 244 373379

Rozga Ante
Faculty of Economics
University of Split
58000 Split
Radovanova 13 CROAZIA
tel.: +385 58 36246
fax: +385 58 366026

Saporta Gilbert
Dept Mathématiques
Conservatoire National des Arts et Métiers
292, Rue Saint-Martin
75141 Paris Cedex 03 FRANCE
tel.: +33 1 40272268
fax: +33 1 40272746
E-mail: saporta@vcnam.cnam.fr

Scarnati Anna
Via D. Frugiuele 34
87100 Cosenza ITALY
tel.: +39 984 71808

Scheaffer Richard L.
Department of Statistics
Griffin-Floyd Hall
University of Florida
102 Griffin-Floyd Hall
Gainesville
Fl. 32611 U.S.A.
tel.: +1 904 3921941
fax: +1 904 3925175
E-mail: scheaffe@stat.ufl.edu

Schuyten Gilberte
Henri Dunantlaan 1
9000 Gent BELGIUM
tel.: +32 9 2646386
fax: +32 9 2646498
E-mail: schuyten@data.rug.ac.be

Scotti Arnaldo
Liceo Scientifico Roiti
Via Mulinetto 81
44100 Ferrara ITALY
tel.: +39 53 27763592

Sernini Marcella
Istituto Tecnico Commerciale
P.F. Calvi
Via Santa Chiara, 10
35123 Padova ITALY
tel.: +39 49 8751966

Severino Liliana
Facultad de Ciencias Economicas
Universidad Nacional de Rosario
Boulevard Oroño 1261
2000 Rosario ARGENTINA
tel.: +54 41 43030
fax: +54 41 43030

Sforza Marta
Universidad Nacional de Rosario
Tarragona 824
2000 Rosario ARGENTINA
fax: +54 41 43030

Sharpe Ken
Dept. of Statistics
University of Melbourne
Parkville 3052
Victoria AUSTRALIA
tel.: +61 3 344 6609
fax: +61 3 344 7499

Shvyrkov Vladislav V.
c/o Statistics Department
Institution National University
536 Oasis Drive
Santa Rosa CA 95407 U.S.A.
fax: +263 4 732828

Smit Chris F.
Dept. of Statistics
University of Pretoria
SOUTH AFRICA
tel.: +27 12 420 3441
fax: +27 12 342 3260
E-mail: sta0005@UPVM2.up.ac.za

Sowey Eric R.
Dept. of Econometrics
University of N.S.W. P.O. Box 1
Kensington, N.S.W. 2033
AUSTRALIA
tel.: +61 2 697 3319
fax: +61 2 313 6337
E-mail: sowey@unsw.edu.au

Sözer E. Emel
Department of Statistics
Faculty of Science
Hacettepe University
06532 Beytepe, Ankara TURKEY
tel.: +90 4 2352500/136
E-mail: 1_EMEL@trhun.bitnet

Starkings Susan
School of Computing, Information
Systems & Mathematics
South Bank University
103 Borough Road
London SE1 0AA U.K.
tel.: +44 71 815 7455
fax: +44 71 815 8155

Stielau Karl
Department of Statistics and Biometry
University of Natal
P.O. Box 375
Pietermaritzburg 3200
SOUTH AFRICA
tel.: +27 331 955508
fax: +27 331 955599
E-mail: stielau@unpsun1.cc.unp.ac.za

Stortini Valter
Istituto Tecnico Commerciale
L. Einaudi
63018 P.to S. Elpidio (AP) ITALY

Stuart Michael
Dept. of Statistics
Trinity College
Dublin 2 IRELAND
E-mail: mstuart@vax1.tcd.ie

Taffe John
Dept. Mathematics
Swinburne University of
Technology
P.O. Box 218
Hawthorn 3122 AUSTRALIA
tel.: +61 3 8198367
fax: +61 3 819 0821
E-mail: jtaffe@buster.cc.swin.edu.au

Teekens Rudolf
Training of European Statisticians
2-6, Route de Trèves - Bloc B/6
L-2633 Senningerberg
LUXEMBOURG
tel.: +352 34 6869-1
fax: +352 34 6873

Tittarelli Luigi
Dip. di Scienze Statistiche
Università di Perugia
Via A. Pascoli
06100 Perugia ITALY
tel.: +39 75 585 5242
fax: +39 75 43242

Valli Marzia
Istituto Statistica Sociale e
Scienze Demografiche e Biometriche
Università di Palermo
Viale delle Scienze
90128 Palermo ITALY
tel.: +39 91 426452

Van Strik Roel
Dept. of Epidemiology and Biost.
Erasmus University Medical School
P.O. Box 1738
3000 DR Rotterdam
NETHERLANDS
tel.: +31 10 4087391
fax: +31 10 4087494

Vännman Kerstin
Department of Mathematics
Lulea University of Technology
951 87 Lulea SWEDEN
fax: +46 9209 8894
E-mail: kerstin.vannman@sm.luth.se

Vass Lucia
Statistics and Legal Informatics
University Eötvös Loránd
Egyetem tér 1-3
H-1364 Budapest HUNGARY
tel.: +36 1 266 5945
fax: +36 1 266 8055

Velev Kamen
Higher Institute of Chemical
Technology
Arch. Milanov 12
1421 Sofia BULGARIA
fax: +359 2 68 33 20

Vere-Jones David
Inst. of Statistics and Op. Research
Victoria University of Wellington
P.O. Box 600
Wellington NEW ZEALAND
tel.: +64 471 5391
fax: +64 495 5118
E-mail: dvj@isor.vuw.ac.nz

Vianelli Bellavista Luciana
Ist. di Matematica per la Ricerca
Operativa
Università di Palermo
Viale delle Scienze
90128 Palermo ITALY
tel.: +39 91 6563231
fax: +39 91 426781

Vitangeli Maria Cesira
Via Statilio Ottato 72 sc.B
00175 Roma ITALY
tel.: +39 6 7610269

Xekalaki Evdokia
Dept. of Statistics
Athens University of Economics and
Business
76 Patision st.
Athens 10434 GREECE
tel.: +30 1 8230488
fax: +30 1 8210509
E-mail: jpan@aueb.ariadne-t.gr

Yang Jingyi
Dept. of Soil Science
Jilin Agricultural University
Changchun Jilin 130118
People's Republic of CHINE

Zuliani Alberto
Dip. Studi Geoeconomici, Statistici e
Storici
Università di Roma "La Sapienza"
Via del Castro Laurenziano 9
00161 Roma ITALY
tel.: +39 6 49766240
fax: +39 6 4957606